Effcient numerical methods for strongly anisotropic elliptic equations

Christophe Besse, Fabrice Deluzet, Claudia Negulescu, Chang Yang

Efficient numerical methods for strongly anisotropic elliptic equations*

Christophe BESSE† Fabrice DELUZET‡ Claudia NEGULESCU§
Chang YANG†¶

March 30, 2011

Abstract

In this paper, we study an efficient numerical scheme for a strongly anisotropic elliptic problem which arises in the modeling of ionospheric plasma dynamics. A small parameter ε induces the anisotropy of the problem, which leads to severe numerical difficulties for $0 < \varepsilon \ll 1$ when solved with standard methods. An AP-scheme is considered in this paper in the 2D resp. 3D physical case with an anisotropy aligned to one coordinate axis and an ε-intensity either constant or variable within the simulation domain. This AP-scheme is uniformly precise in ε, permitting thus the choice of coarse discretization grids, independent of the parameter ε.

Keywords: Anisotropic elliptic equation, Singular Perturbation Model, Asymptotic Preserving scheme, Scharfetter-Gummel scheme, heterogeneous anisotropy ratios.

1 Introduction

Mathematical study and numerical simulation of highly anisotropic problems are among the most delicate tasks in todays computational science, arising in several fields of

*This work has been supported by the Agence Nationale de la Recherche (ANR) under contract IODISEE (ANR-09-COSI-007-02). The first two authors would like to express their gratitude to G. Gallice and C. Tessieras from CEA-Cesta for bringing their attention to this problem.
All the authors would like to acknowledge Pierre Degond for fruitful discussions and his help for improving this paper.

†Laboratoire Paul Painlevé (UMR 8524), Université Lille 1, cité scientifique, 59655 Villeneuve d'Ascq Cedex, France. (christophe.besse@math.univ-lille1.fr, chang.yang@math.univ-lille1.fr)

‡Université de Toulouse, UPS, INSA, UT1, UTM, Institut de Mathématiques de Toulouse, F-31062 Toulouse, France; CNRS, Institut de Mathématiques de Toulouse UMR 5219, F-31062 Toulouse, France. (fabrice.deluzet@math.univ-toulouse.fr)

§CMI/LATP (UMR 6632), université de Provence; 39, rue Joliot Curie; 13453 Marseille Cedes, France. (claudia.negulescu@cmi.univ-mrs.fr)

¶Corresponding author

applications. Examples are flows in porious media [1, 17], semiconductor modeling [25], quasi-neutral plasma simulations [8], image processing [36, 38], atmospheric or oceanic flows [33], the list of possible applications being not exhaustive. The motivation of this work is closely related to the magnetized plasma simulations such as atmospheric plasmas [14, 20, 23, 24, 27].

The difficulties encountered when trying to solve numerically such problems, come from the severe anisotropy of these models, which requires the use of grids refined along the anisotropy direction, being definitively too expensive for real physical simulations. The aim of the present paper is to introduce an efficient numerical scheme for an accurate resolution of such type of problems.

A simplified model can be stated as the following singular perturbation problem, also studied in [10],

$$\begin{cases} -\nabla \cdot (\mathcal{A}\nabla\phi) = f, & \text{in } \Omega, \\ \phi = 0, \text{ on } \partial\Omega_D, \qquad \partial_z\phi = 0, \text{ on } \partial\Omega_z, \end{cases} \tag{1.1}$$

where $\Omega \subset \mathbb{R}^2$ or $\Omega \subset \mathbb{R}^3$ is a rectangular or cuboid domain with boundary $\partial\Omega = \partial\Omega_D \cup \partial\Omega_z$ and $\mathcal{A}$ a diffusion matrix of the form

$$\mathcal{A} = \begin{pmatrix} A_\perp & 0 \\ 0 & \frac{1}{\varepsilon}A_z \end{pmatrix}. \tag{1.2}$$

The terms $A_\perp$ and A_z are of the same order of magnitude, while the parameter $0 < \varepsilon < 1$ may be very small, thus provoking the anisotropy of the problem. This anisotropy is considered along the z-direction, and, as in [10], constant ε parameters are first considered. If we let ε tend to 0, the degenerated equation corresponding to (1.1) is ill-posed when associated with Neumann boundary conditions on $\partial\Omega_z$. This is the reason why standard methods, discretizing the singular perturbation problem (1.1), are not adapted for computations with $\varepsilon \ll 1$, needing a grid mesh dependent on ε, in order to get accurate results. This leads necessarily to high numerical costs and memory usage as well as limited anisotropy ratios. However the limit of the perturbation problem solution is well defined and can be computed by the limit problem derived in section 2.1.

To face this problem, an asymptotic preserving scheme was introduced in [10]. Initially, AP schemes were introduced by S. Jin [21] for the study of multiscale kinetic equations. The main idea of AP-schemes derivation rely on a reformulation of the singular perturbation (SP) problem into an equivalent set of equations for which the limit is regular. With this aim, the solution is decomposed into a mean value, accordingly to the anisotropy ($z-$) direction, corrected by a fluctuation. The mean part of the solution is actually the limit of the (SP)-model solution, the fluctuation being a correction, of magnitude ε, with a zero mean value along the z-coordinate. The discretized reformulated system, referred to as the AP-scheme, was proved in [10] to be efficient to solve the 2D anisotropic elliptic problem (1.1), regardless to ε-values.

In the present paper, we want to introduce some essential improvements of the initial AP scheme. The objective of these improvements and developments is, on the one hand, to enhance the numerical efficiency of the method, and on the other hand, to extend this scheme to a more realistic physical problem, with a particular focus on ionospheric plasma simulations.

The new formulation still relies on a decomposition of the solution in a fluctuating and a mean part, however, in this new formulation, the system verified by the fluctuation is different and much sparser than that of the previous approach, increasing thus considerably the numerical efficiency (reduced computational time and memory usage), making possible to address three dimensional problems with limited resources, as shall be demonstrated in this paper. Moreover, two strategies are proposed and compared in the present paper for the resolution of the obtained sparse linear system. The two components of the solution (*i.e.*, the mean and fluctuating part) verify a system of coupled equations. The iterative method initially proposed in [10] for the resolution of this coupled system is compared with a direct resolution for both components. These two new features of the here proposed AP-reformulation, *i.e.* the sparser linear system and the direct resolution of the coupled system will induce considerable numerical savings and make this AP scheme much more performant for real physical simulations as compared to the original AP scheme introduced in [10]. The capabilities of this new AP-scheme will give rise to a the forthcoming paper [6] addressing real ionospheric plasma simulation.

Furthermore, we are interested in the present paper in non homogeneous anisotropy ratios accordingly to real physical configurations. For some applications (for instance ionospheric plasma physic) large variations and steep gradients can be observed in this ratio. This feature has motivated the development of this new AP-scheme to handle highly heterogeneous anisotropies. To improve the accuracy of the scheme in this framework, a Scharfetter-Gummel [31] version of the method is proposed and compared to standard formulations.

Finally, a 3D extension of the AP-model is introduced. The diffusion block corresponding to the perpendicular direction $A_\perp$ is no more diagonal, but can contain variable non-diagonal terms coming from the transverse mobilities of the 3-dimensional ionospheric plasma model [4]. We develop further the AP scheme to this case and demonstrate its efficiency by numerical simulations with constant ε as well as a variable ε-function.

Let us remark that a different approach is presented in [11] in the aim to solve similar anisotropic elliptic problems. The AP reformulation proposed there is based on a different decomposition of the unknown function ϕ (micro-macro decomposition), which permits to avoid the introduction of Lagrange multipliers, essential in the here proposed AP version, to take into account for the fluctuation part zero mean constraint (see section 2). However, the AP reformulation of [11] does not take into account for the high anisotropy gradients as well as for possible non diagonal terms in the $A_\perp$ submatrix. Moreover the method here introduced, is derived as a correction of any existing code providing the solution of the original anisotropic elliptic equation (1.1). As demonstrated in section 2.5.3, the AP scheme is obtained by means of limited enhancements of the original singular perturbation problem. These points are crucial for a real physical simulation and shall be investigated here.

This paper is organized as follows: in section 2, we recall the AP-scheme of [10], based on an iterative resolution, and introduce the new AP-version as well as its weak formulation. The existence and uniqueness of the reformulated continuous and discrete systems solution as well as the convergence of the iterative resolution are demonstrated. The two approaches are then discretized by a finite element method of type $\mathbb{Q}_1$ and the

numerical results are compared for values of ε below the machine aritmethic precision. In section 3, the AP-scheme is extended for the variable ε-case, with large gradients. Discretizations and numerical results are presented. Finally, in section 4, we study a 3D anisotropic elliptic problem which has the same structure as a real 3D ionospheric model, containing non-diagonal terms in the perpendicular matrix-block $A_\perp$. The numerical results of the constant ε case and the variable ε case are compared, considering different mesh sizes.

2 Derivation of a new asymptotic preserving scheme for a uniform anisotropy ratio ε

In this section, a new Asymptotic Preserving reformulation is introduced for the singularly perturbed problem (1.1) in a framework comparable to that of the previous study [10], *i.e.* we are considering the constant ε case. The advantages of this new AP-scheme as compared with the original one presented in [10] shall be outlined here.

The present section is organized as follows. Some common properties of the singular perturbation system are firstly recalled in subsection 2.1. The new AP formulation is then stated in subsection 2.2. Subsections 2.3 and 2.4 concern some mathematical investigations (existence/uniqueness of solutions). The numerical discretization is briefly detailed in subsection 2.5, before carrying out the numerical experiments .

2.1 Investigation of the singular perturbation model and its limit regime

For simplicity reasons, we first consider a 2-dimensional model with Ω a rectangular domain defined as $\Omega = \Omega_x \times \Omega_z$ where $\Omega_x \subset \mathbb{R}$ and $\Omega_z \subset \mathbb{R}$ are intervals. The coefficients $A_\perp$ and A_z of the diffusion matrix $\mathcal{A}$ are scalar functions in this case. A 3-dimensional case will be considered in the last part of this paper.

The 2-dimensional elliptic problem is given as

$$\begin{cases} -\nabla \cdot (\mathcal{A}\nabla\phi^\varepsilon) = f, & \text{in } \Omega, \\ \phi^\varepsilon = 0, \text{ on } \partial\Omega_x \times \Omega_z, \quad \partial_z\phi^\varepsilon = 0, \text{ on } \Omega_x \times \partial\Omega_z, \end{cases} \tag{2.1}$$

where $\mathcal{A}$ is given by (1.2) and $A_\perp(x,z)$ and $A_z(x,z)$ are known functions of the same order of magnitude. The singularly perturbed problem (2.1) can be rewritten as (SP-model)

$$(\mathrm{SP}) \begin{cases} -\dfrac{\partial}{\partial x}\left(\varepsilon A_\perp \dfrac{\partial\phi^\varepsilon}{\partial x}\right) - \dfrac{\partial}{\partial z}\left(A_z \dfrac{\partial\phi^\varepsilon}{\partial z}\right) = \varepsilon f, & \text{in } \Omega, \\ \phi^\varepsilon = 0, & \text{on } \partial\Omega_x \times \Omega_z, \\ \partial_z\phi^\varepsilon = 0, & \text{on } \Omega_x \times \partial\Omega_z. \end{cases} \tag{2.2}$$

In the rest of this paper, we shall suppose the following hypothesis

Hypothesis 1. *Let the diffusion functions $A_\perp \in L^\infty(\Omega)$ and $A_z \in L^\infty(\Omega)$ satisfy $0 < c_\perp \leq A_\perp(x,z) \leq M_\perp$, $0 < c_z \leq A_z(x,z) \leq M_z$, where $c_\perp$, c_z, $M_\perp$, M_z are some positive constants. Moreover let $f \in L^2(\Omega)$.*

The model (2.2) is a well-posed boundary value problem which has a unique solution for all fixed $\varepsilon > 0$. However setting formally $\varepsilon = 0$ in this equation, we obtain a degenerate problem reading:

$$\begin{cases} -\dfrac{\partial}{\partial z}\left(A_z \dfrac{\partial \psi}{\partial z}\right) = 0, & \text{in } \Omega, \\ \psi = 0, & \text{on } \partial\Omega_x \times \Omega_z, \\ \partial_z \psi = 0, & \text{on } \Omega_x \times \partial\Omega_z. \end{cases} \tag{2.3}$$

The model (2.3) is ill-posed due to the loss of uniqueness. Indeed, all the functions ψ depending only on the x-coordinate and verifying the boundary condition $\psi = 0$ on $\partial\Omega_x$ satisfy (2.3). However, ϕ^0, the $\varepsilon \to 0$ limit of the sequence ϕ^ε consisting of the solutions of the singular perturbation problem (2.2), is unique and can be determined by a well posed problem called in the sequel L-problem (Limit problem). To get this limit problem, we first integrate the singular perturbation problem along the anisotropy direction, which yields

$$-\frac{\partial}{\partial x}\left(\overline{A_\perp \frac{\partial \phi^\varepsilon}{\partial x}}\right) = \overline{f}, \tag{2.4}$$

where $\overline{f}$ denotes the mean value of the function f along the anisotropy direction, *i.e.*

$$\overline{f}(x) = \frac{1}{\mathrm{mes}(\Omega_z)} \int_{\Omega_z} f(x,z)\, dz\,, \qquad \forall x \in \Omega_x\,.$$

Secondly, the limit ϕ^0 will also verify (2.3), and will be thus independent of the z-coordinate. Passing to the limit in (2.4) and using this property provides the Limit problem (L-model), stated as: Find $\phi^0 : \Omega_x \to \mathbb{R}$ solution of

$$(\mathrm{L}) \begin{cases} -\dfrac{\partial}{\partial x}\left(\overline{A}_\perp \dfrac{\partial \phi^0}{\partial x}\right) = \overline{f}, & \text{in } \Omega_x\,, \\ \phi^0 = 0\,, & \text{on } \partial\Omega_x\,. \end{cases} \tag{2.5}$$

Having now introduced the SP-problem and its limit problem (L), the aim is to introduce an equivalent reformulation of the SP-problem, which shall permit to pass continuously from the SP-problem (2.2) to the Limit-problem (2.5) as $\varepsilon \to 0$. The SP-formulation (2.2) does not "verify" this feature, as we mentioned above, however this is the main characteristics of the class of AP-schemes.

The main idea of the Asymptotic Preserving method introduced in [10] is to decompose the singular perturbation solution ϕ^ε into two parts: $\bar{\phi}^\varepsilon$, the mean part integrated along the z-axis, complemented with the fluctuating part $\phi^{\varepsilon\prime}$ defined by $\phi^{\varepsilon\prime} = \phi^\varepsilon - \bar{\phi}^\varepsilon$. With these new unknowns, the system (2.2) is decomposed into respectively an average equation for $\bar{\phi}^\varepsilon$ and a fluctuation equation for $\phi^{\varepsilon\prime}$. The average equation is very similar to the limit problem (2.5) and provides an accurate manner for computing the mean part for all $\varepsilon > 0$. Moreover, the fluctuating part verifies $\overline{\phi^{\varepsilon\prime}} = 0$, which is the fundamental property in deriving the AP-scheme, and in particular which shall permit to get an accuracy independent of the parameter ε. This point is detailed in subsection 2.2.

2.2 Asymptotic preserving formulation

For the sake of simplicity, let $\Omega_x = [0, L_x]$ and $\Omega_z = [0, L_z]$ and let us omit the ε-index of the solution ϕ^ε whenever there is no confusion. There are some useful properties of the average and fluctuation operations, listed below

$$\overline{f'} = 0,\ \overline{fg} = \bar{f}\bar{g} + \overline{f'g'},\ \overline{f'g'} = \overline{f'g} = \overline{fg'}, \tag{2.6}$$

$$\overline{\frac{\partial f}{\partial x}} = \frac{\partial \bar{f}}{\partial x},\ \frac{\partial f}{\partial z} = \frac{\partial f'}{\partial z},\ \left(\frac{\partial f}{\partial x}\right)' = \frac{\partial f'}{\partial x}, \tag{2.7}$$

$$(fg)' = f'g' - \overline{f'g'} + \bar{f}g' + f'\bar{g}. \tag{2.8}$$

Taking now the average of the equation (2.2) along the z-axis and considering the properties (2.6), the equation for the mean part of the solution states:

$$\begin{cases} -\dfrac{\partial}{\partial x}\left(\overline{A_\perp}\dfrac{\partial \bar{\phi}}{\partial x}\right) = \bar{f} + \dfrac{\partial}{\partial x}\left(\overline{A_\perp \dfrac{\partial \phi'}{\partial x}}\right), & \text{in } \Omega_x, \\ \bar{\phi} = 0, & \text{on } \partial\Omega_x. \end{cases} \tag{2.9}$$

The equation for the fluctuating part, proposed in this paper, is straightforwardly derived by introducing the decomposition $\phi(x,z) = \phi'(x,z) + \bar{\phi}(x)$ in (2.2) and applying (2.7), yielding thus

$$\begin{cases} -\varepsilon\dfrac{\partial}{\partial x}\left(A_\perp\dfrac{\partial \phi'}{\partial x}\right) - \dfrac{\partial}{\partial z}\left(A_z\dfrac{\partial \phi'}{\partial z}\right) = \varepsilon f + \varepsilon\dfrac{\partial}{\partial x}\left(A_\perp\dfrac{\partial \bar{\phi}}{\partial x}\right), & \text{in } \Omega, \\ \phi' = 0, & \text{on } \partial\Omega_x \times \Omega_z, \\ \partial_z \phi' = 0, & \text{on } \Omega_x \times \partial\Omega_z, \\ \overline{\phi'} = 0, & \text{in } \Omega_x. \end{cases} \tag{2.10}$$

The system (2.9),(2.10) will be referred to as the New AP-formulation. It differs from the one introduced in the original AP-paper [10] by the equation for the fluctuating part, whose derivation is briefly detailed here for comparison purpose. Subtracting the average equation (2.9) from (2.2) and using (2.8), one gets the different fluctuation system

$$\begin{cases} -\varepsilon\dfrac{\partial}{\partial x}\left(A_\perp\dfrac{\partial \phi'}{\partial x}\right) - \dfrac{\partial}{\partial z}\left(A_z\dfrac{\partial \phi'}{\partial z}\right) + \varepsilon\dfrac{\partial}{\partial x}\left(\overline{A'_\perp\dfrac{\partial \phi'}{\partial x}}\right) = \\ \qquad\qquad \varepsilon f' + \varepsilon\dfrac{\partial}{\partial x}\left(A'_\perp\dfrac{\partial \bar{\phi}}{\partial x}\right), & \text{in } \Omega, \\ \phi' = 0, & \text{on } \partial\Omega_x \times \Omega_z, \\ \partial_z \phi' = 0, & \text{on } \Omega_x \times \partial\Omega_z, \\ \overline{\phi'} = 0, & \text{in } \Omega_x. \end{cases} \tag{2.11}$$

The coupled system (2.9),(2.11) was introduced and analyzed in [10]. The two reformulation (2.9),(2.10) resp. (2.9),(2.11) are equivalent. We investigate below their asymptotic

preserving properties and their ability to provide a precise computation of the solution of (2.2) for all values of ε.

First, note that the average equation (2.9) is a well-posed boundary value problem, which is independent of ε. Moreover, letting $\varepsilon \to 0$ in the fluctuation equation (2.10) or (2.11), yields

$$\begin{cases} -\dfrac{\partial}{\partial z}\left(A_z \dfrac{\partial \phi'}{\partial z}\right) = 0, & \text{in } \Omega, \\ \phi' = 0, & \text{on } \partial\Omega_x \times \Omega_z, \\ \partial_z \phi' = 0, & \text{on } \Omega_x \times \partial\Omega_z, \\ \overline{\phi'} = 0, & \text{in } \Omega_x. \end{cases} \tag{2.12}$$

In contrast to (2.3), this problem is well-posed with unique solution $\phi' \equiv 0$. Indeed, it is the constraint $\overline{\phi'} = 0$ which implies the uniqueness of the solution. And it is also this property of zero mean value for the fluctuating part ϕ' in (2.10) resp. (2.11), which provides an equation with a condition number independent of the ε-values.

Now, setting $\phi' = 0$ into the average equation (2.9), yields the limit model (2.5). This demonstrates that the reformulated systems (2.9),(2.10) resp. (2.9),(2.11) are regular perturbations of the L-problem for $\varepsilon \to 0$.

The equation (2.11) for the fluctuating part ϕ' has been designed in [10] in order to have a zero mean value right hand-side, thus ensuring that the fluctuating part itself verifies this property and justifying by this manner the introduction of the crucial constraint $\overline{\phi'} = 0$. However this equation incorporates a term, namely $\frac{\partial}{\partial x}\left(\overline{A'_\perp \frac{\partial \phi'}{\partial x}}\right)$, which fills the matrix in the discretization step. For this reason, this paper will focus on the sparse alternative (2.9),(2.10), more efficient in terms of numerical memory usage and computations. Due to the equivalence of the two AP-reformulations, one has the important property $\bar{\phi}' \equiv 0$ even in system (2.9),(2.10), although the right hand-side of (2.10) does not verify this zero mean value property.

Note that for both formulations, the equations providing the mean and the fluctuating parts are coupled. Two strategies will be proposed in subsection 2.4 in order to solve this coupled system.

2.3 Weak formulation

In order to introduce the weak form of the AP-system (2.9),(2.10), let us introduce two Hilbert-spaces

$$\mathbb{V} = \{\psi(x,z) \in H^1(\Omega)/\psi = 0 \text{ on } \partial\Omega_x \times \Omega_z\}, \ \mathbb{W} = \{\bar{\psi}(x) \in H^1(\Omega_x)/\bar{\psi} = 0 \text{ on } \partial\Omega_x\},$$

and the corresponding scalar products

$$(\phi,\psi)_{\mathbb{V}} = \varepsilon(\partial_x\phi, \partial_x\psi)_{L^2(\Omega)} + (\partial_z\phi, \partial_z\psi)_{L^2(\Omega)}, \ (\bar{\phi},\bar{\psi})_{\mathbb{W}} = (\partial_x\bar{\phi}, \partial_x\bar{\psi})_{L^2(\Omega_x)}.$$

For simplicity, we denote in the sequel the L^2 scalar-product simply by the bracket $(\cdot,\cdot)$. By defining the following bilinear forms

$$\begin{aligned}
a_0(\phi',\psi') &:= \int_0^{Lz}\int_0^{Lx} A_z(x,z)\frac{\partial\phi'}{\partial z}(x,z)\frac{\partial\psi'}{\partial z}(x,z)dxdz,\\
a_1(\phi',\psi') &:= \int_0^{Lz}\int_0^{Lx} A_\perp(x,z)\frac{\partial\phi'}{\partial x}(x,z)\frac{\partial\psi'}{\partial x}(x,z)dxdz,\\
a_2(\bar\phi,\bar\psi) &:= \int_0^{Lx} \bar A_\perp(x)\frac{\partial\bar\phi}{\partial x}(x)\frac{\partial\bar\psi}{\partial x}(x)dx,\\
a(\phi',\psi') &:= a_0(\phi',\psi') + \varepsilon a_1(\phi',\psi'),\\
b(\bar P,\psi') &:= \int_0^{Lx}\bar P(x)\int_0^{Lz}\psi'(x,z)dzdx,\\
c(\bar\phi,\psi') &:= \int_0^{Lz}\int_0^{Lx} A_\perp(x,z)\frac{\partial\bar\phi}{\partial x}(x)\frac{\partial\psi'}{\partial x}(x,z)dxdz,
\end{aligned} \tag{2.13}$$

we can write the weak formulations of the SP-problem resp. L-problem as follows

$$\text{(SP)}\qquad \varepsilon a_1(\phi^\varepsilon,\psi)+a_0(\phi^\varepsilon,\psi)=\varepsilon(f,\psi),\ \forall\psi\in\mathbb{V}, \tag{2.14}$$

$$\text{(L)}\qquad a_2(\phi^0,\psi^0)=(\bar f,\psi^0),\ \forall\psi^0\in\mathbb{W}. \tag{2.15}$$

Introducing in (SP) the decomposition $\phi^\varepsilon=\bar\phi^\varepsilon+\phi^{\varepsilon\prime}$ and taking test-function $\psi'\in\mathbb{V}$ resp. $\bar\psi\in\mathbb{W}$ gives rise to the following equivalent reformulation of the SP-problem

$$\begin{cases} a_2(\bar\phi,\bar\psi)=(\bar f,\bar\psi)-\frac{1}{L_z}c(\bar\psi,\phi'), & \forall\bar\psi\in\mathbb{W},\\ a(\phi',\psi')=\varepsilon(f,\psi')-\varepsilon c(\bar\phi,\psi'), & \forall\psi'\in\mathbb{V}. \end{cases} \tag{2.16}$$

In order to remain well-posed even in the limit of $\varepsilon\to 0$, we have to introduce the constraint $\overline{\phi'}\equiv 0$ into the fluctuation equation as mentioned in subsection 2.2. This is realized via the introduction of a Lagrange multiplier $\bar P$ [10] as follows : Find $(\overline{\phi},\phi',\overline{P})\in\mathbb{W}\times\mathbb{V}\times L^2(\Omega_x)$, solution of

$$(AP)\begin{cases} a_2(\bar\phi,\bar\psi)=(\bar f,\bar\psi)-\frac{1}{L_z}c(\bar\psi,\phi'), & \forall\bar\psi\in\mathbb{W}, \quad (a)\\ a(\phi',\psi')+b(\bar P,\psi')=\varepsilon(f,\psi')-\varepsilon c(\bar\phi,\psi'), & \forall\psi'\in\mathbb{V}, \quad (b)\\ b(\bar Q,\phi')=0, & \forall\bar Q\in L^2(\Omega_x). \quad (c) \end{cases} \tag{2.17}$$

This system will be called in the sequel the AP-reformulation of the SP-problem. The next theorem proves the well-posedness of the AP-formulation (2.17) for fixed $\varepsilon>0$, the equivalence with problem (2.16) and analyzes the convergence of the solution $(\bar\phi^\varepsilon,\phi^{\varepsilon\prime})\in\mathbb{W}\times\mathbb{V}$. In particular it shows also that the AP-problem is well-posed even in the limit $\varepsilon\to 0$.

Theorem 2.1. *Under hypothesis 1, for all fixed $\varepsilon>0$, there exists a unique solution $(\bar\phi^\varepsilon,\phi^{\varepsilon\prime},\bar P^\varepsilon)\in\mathbb{W}\times\mathbb{V}\times L^2(\Omega_x)$ satisfying (2.17). The function $\phi^\varepsilon=\bar\phi^\varepsilon+\phi^{\varepsilon\prime}$ is the unique solution of (2.14). Furthermore, the two systems (2.16) and (2.17) are equivalent. In particular $(\bar\phi^\varepsilon,\phi^{\varepsilon\prime})\in\mathbb{W}\times\mathbb{V}$ is the unique solution of (2.16) if and only if $(\bar\phi^\varepsilon,\phi^{\varepsilon\prime},\bar P^\varepsilon)\in$*

$\mathbb{W} \times \mathbb{V} \times L^2(\Omega_x)$ *is the unique solution of (2.17), with* $\bar{P}^\varepsilon \equiv 0$. *Finally, in the limit* $\varepsilon \to 0$, *the pair* $(\bar{\phi}^\varepsilon, \phi^{\varepsilon\prime})$ *converges towards some function* $(\bar{\phi}^0, \phi^{0\prime}) \in \mathbb{W} \times \mathbb{V}$, *where* $\bar{\phi}^0$ *is the unique weak solution of the L-problem and* $\phi^{0\prime} \equiv 0$.

Proof. This theorem was proven in [10] for the system (2.9),(2.11). Using the equivalence between the AP-formulation of [10] (2.9),(2.11) and the present one (2.9),(2.10), one can immediately adapt the proof.

□

Remark that if we ask more regularity of the coefficients $A_\perp$ and A_z, then we can obtain a more regular solution of the SP-model (2.2). This is a simple regularity result for elliptic equations.

Hypothesis 2. *Let the diffusion functions* $A_\perp \in W^{1,\infty}(\Omega)$ *and* $A_z \in W^{1,\infty}(\Omega)$ *satisfy* $0 < c_\perp \leq A_\perp(x,z) \leq M_\perp$, $0 < c_z \leq A_z(x,z) \leq M_z$, *where* $c_\perp$, c_z, $M_\perp$, M_z *are some positive constants. Moreover let* $f \in L^2(\Omega)$.

Lemma 2.2. *Under hypothesis 2, for all fixed* $\varepsilon > 0$, *there exists a unique solution of the SP-problem (2.2), satisfying* $\phi \in \mathbb{V} \cap H^2(\Omega)$. *Furthermore, we can develop the same arguments as in Theorem 2.1 by replacing* $\mathbb{V}$ *resp.* $\mathbb{W}$ *by* $\mathbb{V} \cap H^2(\Omega)$ *resp.* $\mathbb{W} \cap H^2(\Omega_x)$.

2.4 Iterative procedure

The coupled AP-system (2.17) can be solved either directly, or iteratively, as proposed in [10]. Let us investigate in this subsection the convergence of the iterative procedure.

Let us for this define the Hilbert space

$$\mathbb{U} := \{\phi \in \mathbb{V} \ / \ \bar{\phi} = 0\},$$

associated with the scalar product

$$(\phi, \psi)_{\mathbb{U}} = \int_{\Omega_x} \int_{\Omega_z} A_z \partial_z \phi \, \partial_z \psi dz \, dx + \varepsilon \int_{\Omega_x} \int_{\Omega_z} A_\perp \partial_x \phi \, \partial_x \psi dz \, dx \, .$$

Then we have

Proposition 2.3. *Let hypothesis 2 be satisfied and let us fix* $\varepsilon > 0$. *Moreover, let us construct a fixed point map* $T : \mathbb{U} \to \mathbb{U}$ *as follows*

$$T : \phi' \in \mathbb{U} \xmapsto{(2.17a)} \bar{\phi} \in \mathbb{W} \xmapsto{(2.17b),(2.17c)} \theta' = T(\phi') \in \mathbb{U}.$$

Then for every starting point $\phi'_0 \in \mathbb{U}$, *the sequence* $\phi'_k := T(\phi'_{k-1}) = T^k(\phi'_0)$ *converges in* $\mathbb{U}$ *towards the unique fixed point* $\phi'_* \in \mathbb{U}$ *of* T.

Proof. We prove first that the application T is well-defined. This part of the proof is more delicate than in [10], such that we shall detail it here. The well-posedness of the first step $T_1 : \phi' \in \mathbb{U} \xmapsto{(2.17a)} \bar{\phi} \in \mathbb{W}$ is immediate by the Lax-Milgram lemma. The second step is equivalent to the question: for some $\bar{\phi} \in \mathbb{W}$ solving (2.17a), is there

a unique $(\theta', \bar{P}) \in \mathbb{U} \times L^2(\Omega_x)$ solving (2.17b), (2.17c)? Let us thus investigate the following saddle-point problem: Find $(\theta', \bar{P}) \in \mathbb{U} \times L^2(\Omega_x)$ solution of

$$\begin{cases} a(\theta', \psi') + b(\bar{P}, \psi') = \varepsilon(f, \psi') - \varepsilon c(\bar{\phi}, \psi'), & \forall \psi' \in \mathbb{V}, \\ b(\bar{Q}, \theta') = 0, & \forall \bar{Q} \in L^2(\Omega_x). \end{cases} \tag{2.18}$$

To solve this problem, let us recall that the two AP-formulations (2.9),(2.10) resp. (2.9),(2.11) are equivalent. Thus for the given $\bar{\phi} \in \mathbb{W}$, solution of (2.17a), let us define $\theta' \in \mathbb{U}$ as the solution of

$$a(\theta', \psi') - \varepsilon \int_\Omega \overline{A'_\perp \frac{\partial \theta'}{\partial x}} \frac{\partial \psi'}{\partial x} dxdz = \varepsilon(f', \psi') + \varepsilon \int_\Omega A'_\perp \frac{\partial \bar{\phi}}{\partial x} \frac{\partial \psi'}{\partial x} dxdz, \forall \psi' \in \mathbb{V}. \tag{2.19}$$

The existence and uniqueness of such a solution was proven in [10]. Defining then $\bar{P}$ as

$$\bar{P}(x) = \varepsilon \frac{\partial}{\partial x} \left(\overline{A'_\perp \frac{\partial}{\partial x}(\theta' - \phi')} \right),$$

one can prove immediately that $(\theta', \bar{P}) \in \mathbb{U} \times L^2(\Omega_x)$ solves (2.18).

The uniqueness of the solution of (2.18) is immediate. Taking two different solutions $(\theta'_1, \bar{P}_1), (\theta'_2, \bar{P}_2) \in \mathbb{U} \times L^2(\Omega_x)$, and denoting $\tilde{\theta}' := \theta'_1 - \theta'_2$, $\tilde{P} := \bar{P}_1 - \bar{P}_2$, we have

$$\begin{cases} a(\tilde{\theta}', \psi') + b(\tilde{P}, \psi') = 0, & \forall \psi' \in \mathbb{V}, \\ b(\bar{Q}, \tilde{\theta}') = 0, & \forall \bar{Q} \in L^2(\Omega_x). \end{cases}$$

Choosing now $\psi' := \tilde{\theta}'$ implies $a(\tilde{\theta}', \tilde{\theta}') = 0$. Thus $\tilde{\theta}' \equiv 0$ by coercivity arguments and $b(\tilde{P}, \psi') = 0, \forall \psi' \in \mathbb{V}$. By density arguments this last equation is even valid for all $\psi' \in L^2(\Omega_x)$, such that taking $\psi' := \tilde{P}$, implies $b(\tilde{P}, \tilde{P}) = 0$, hence $\tilde{P} \equiv 0$. As a result, there is a unique $(\theta', \bar{P}) \in \mathbb{U} \times L^2(\Omega_x)$ solving (2.17b), (2.17c) and the fixed point map T is thus well-defined.

The next step is to prove that T is contractive, i.e. for $\varphi'_1, \varphi'_2 \in \mathbb{U}$ to show that

$$||T(\varphi'_1) - T(\varphi'_2)||_{\mathbb{U}} < ||\varphi'_1 - \varphi'_2||_{\mathbb{U}}.$$

However, this part of the proof remains the same as in [10], so that we refer the reader to this previous work. □

2.5 Numerical methods and experiments

The aim of this section shall be to introduce a numerical discretization of the AP-formulation (2.17), to investigate the existence and uniqueness of discrete resolutions and to analyze the obtained results.

2.5.1 A finite element discretization

To discretize system (2.17), we introduce the homogeneous partitions, *i.e.* $x_i = i\Delta x$, $i = 0, \dots, N_x + 1$, and $z_k = k\Delta z$, $k = 0, \dots, N_z + 1$, and the finite element $\mathbb{P}_1$ hat functions

$$\chi_i(x) = \begin{cases} \frac{x - x_{i-1}}{\Delta x}, & x \in [x_{i-1}, x_i), \\ \frac{x_{i+1} - x}{\Delta x}, & x \in [x_i, x_{i+1}), \\ 0, & \text{else}, \end{cases} \qquad i = 1, \cdots, N_x,$$

$$\kappa_0(z) = \begin{cases} \frac{z_1 - z}{\Delta z}, & z \in [z_0, z_1), \\ 0, & \text{else}, \end{cases}$$

$$\kappa_k(z) = \begin{cases} \frac{z - z_{k-1}}{\Delta z}, & z \in [z_{k-1}, z_k), \\ \frac{z_{k+1} - z}{\Delta z}, & z \in [z_k, z_{k+1}), \\ 0, & \text{else}, \end{cases} \qquad k = 1, \cdots, N_z,$$

$$\kappa_{N_z+1}(z) = \begin{cases} \frac{z - z_{N_z}}{\Delta z}, & z \in [z_{N_z}, z_{N_z+1}), \\ 0, & \text{else}. \end{cases}$$

Note that in a Cartesian mesh the tensor product of $\mathbb{P}_1$ hat functions χ_i and κ_k coincides with $\mathbb{Q}_1$ finite element basis functions. The discrete spaces $\mathbb{V}_h \subset \mathbb{V}$, $\mathbb{W}_h \subset \mathbb{W}$ and $\mathbb{L}_h \subset L^2(\Omega_x)$ are generated respectively via the basis functions $(\chi_i)_{i=1,\dots,N_x}$ and $(\kappa_k)_{k=0,\dots,N_z+1}$. Thus we can express the approximations of the unknowns $\phi'_h \in \mathbb{V}_h$, $\bar{\phi}_h \in \mathbb{W}_h$, $\bar{P}_h \in \mathbb{L}_h$ in the form

$$\phi'_h(x,z) = \sum_{i=1}^{N_x} \sum_{k=0}^{N_z+1} \alpha_{ik} \chi_i(x) \kappa_k(z), \;\; \bar{\phi}_h(x) = \sum_{i=1}^{N_x} \beta_i \chi_i(x), \;\; \bar{P}_h(x) = \sum_{i=1}^{N_x} \gamma_i \chi_i(x). \tag{2.20}$$

The discretized AP-problem can now be expressed as follows: Find $(\overline{\phi_h}, \phi'_h, \overline{P}_h) \in \mathbb{W}_h \times \mathbb{V}_h \times \mathbb{L}_h$ solution of

$$(AP)_h \begin{cases} a_2(\bar{\phi}_h, \bar{\psi}_h) = (\bar{f}, \bar{\psi}_h) - \frac{1}{L_z} c(\bar{\psi}_h, \phi'_h), & \forall \bar{\psi}_h \in \mathbb{W}_h, \\ a(\phi'_h, \psi'_h) + b(\bar{P}_h, \psi'_h) = \varepsilon(f, \psi'_h) - \varepsilon c(\bar{\phi}_h, \psi'_h), & \forall \psi'_h \in \mathbb{V}_h, \\ b(\bar{Q}_h, \phi'_h) = 0, & \forall \bar{Q}_h \in \mathbb{L}_h. \end{cases} \tag{2.21}$$

Denoting by $A_2 \in \mathbb{R}^{N_x \times N_x}$, $A \in \mathbb{R}^{N_x(N_z+2) \times N_x(N_z+2)}$, $B, C \in \mathbb{R}^{N_x \times N_x(N_z+2)}$ the matrices associated with the bilinear forms a_2, a, b, c respectively and moreover let us define the right-hand sides $F_1 \in \mathbb{R}^{N_x(N_z+2)}$, $F_2 \in \mathbb{R}^{N_x}$ by

$$F_1(ik) := \varepsilon(f, \chi_i \kappa_k), \quad F_2(i) := \varepsilon(\bar{f}, \chi_i), \qquad \forall i = 1, \dots, N_x, \; k = 0, \dots, N_z + 1,$$

then the discrete system can then be recasted in : Find $(\alpha, \beta, \gamma) \in \mathbb{R}^{N_x(N_z+2)} \times \mathbb{R}^{N_x} \times \mathbb{R}^{N_x}$ solution of

$$\begin{pmatrix} A & \varepsilon C & B \\ \varepsilon C^T & \varepsilon A_2 & 0 \\ B^T & 0 & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} F_1 \\ F_2 \\ 0 \end{pmatrix}. \tag{2.22}$$

Remark that we multiplied the second equation (2.21) by εL_z in order to get a symmetric matrix.

Thanks to the theorem 2.1 and the proposition 2.3, two strategies are proposed for the resolution of this linear system, a direct resolution consisting in solving directly (2.22) with an appropriate linear solver and an iterative resolution based on the following decoupling

$$\varepsilon A_2 \beta^{(n+1)} = F_2 - \varepsilon C^T \alpha^{(n)}\,, \tag{2.23a}$$

$$\begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} \alpha^{(n+1)} \\ \gamma^{(n+1)} \end{pmatrix} = \begin{pmatrix} F_1 \\ 0 \end{pmatrix} - \begin{pmatrix} 0 & \varepsilon C \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ \beta^{(n+1)} \end{pmatrix}. \tag{2.23b}$$

It consists in solving the equation for the mean part using an estimation of the fluctuating part, and then to compute a better estimate of the fluctuating part thanks to the updated mean approximation. In the next subsection, we shall prove the existence and uniqueness of a discrete solution corresponding to these two approaches.

2.5.2 Existence/Uniqueness of a direct resp. iterative resolution

We have seen in the previous subsections that the AP-formulation (2.17) has a unique weak solution $(\overline{\phi}, \phi', \overline{P}) \in \mathbb{W} \times \mathbb{V} \times L^2(\Omega_x)$. The aim of the present subsection is to prove that the discrete AP-formulation (2.21) has also a unique solution, in other words, that the linear system (2.22) is invertible.

Theorem 2.4. *Let $\varepsilon > 0$ be fixed and let us suppose that hypothesis 1 is satisfied. Then the discrete AP-formulation (2.21) admits a unique solution $(\overline{\phi_h}, \phi'_h, \overline{P}_h) \in \mathbb{W}_h \times \mathbb{V}_h \times \mathbb{L}_h$.*

Proof. As we consider now a finite dimensional linear system, we have only to check the uniqueness of the solution. For this, let $f \equiv 0$ and let us show that this implies $(\overline{\phi_h}, \phi'_h, \overline{P}_h) \equiv 0$. The proof is very similar to the proof of the uniqueness of the solution in the continuous case.

The first step is to show that $\overline{P}_h = 0$. For this let us take in the second equation of (2.21) test functions $\psi'_h \in \mathbb{W}_h$ depending only on the x-coordinate. Using the first equation, this implies $b(\bar{P}_h, \psi'_h) = 0$ for all $\psi'_h \in \mathbb{W}_h$, yielding immediately $\overline{P}_h = 0$.
The second step is to show that $(\overline{\phi_h}, \phi'_h) \equiv (0, 0)$. For this, let us take in the second equation of (2.21) as test function $\psi'_h := \overline{\phi_h} + \phi'_h \in \mathbb{V}_h$. This yields

$$\int_0^{L_x} \int_0^{L_z} A_z |\partial_z \phi'_h|^2 dz dx + \varepsilon \int_0^{L_x} \int_0^{L_z} A_\perp |\partial_x (\phi'_h + \overline{\phi_h})|^2 dz dx = 0\,,$$

implying thus $\overline{\phi_h} + \phi'_h \equiv cst$. Due to the Dirichlet boundary conditions, one gets $\overline{\phi_h} + \phi'_h \equiv 0$. The third equation of (2.21) however states (after some simple computations) that $\bar{\phi}'_h = 0$, such that integrating in z the equation $\overline{\phi_h} + \phi'_h \equiv 0$ yields immediately $(\overline{\phi_h}, \phi'_h) \equiv (0, 0)$ and we have finished the proof. □

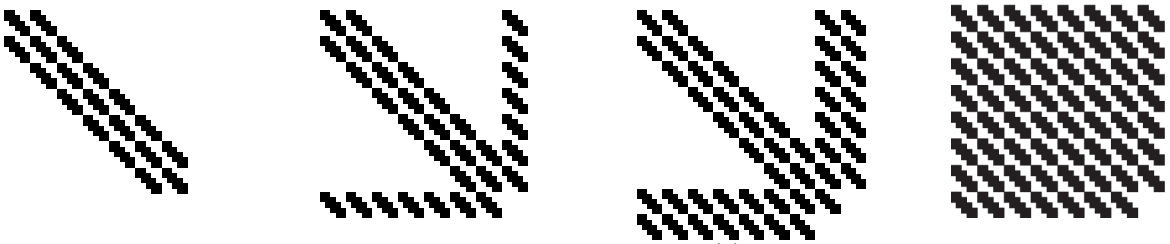
As we did for the direct resolution, we can prove now that the iterative procedure has also a unique solution.

Proposition 2.5. *Let $\varepsilon > 0$ be fixed and let us suppose that hypothesis 2 is satisfied. Then the iterative procedure of the AP-formulation (2.23) admits a unique solution $(\overline{\phi_h}, \phi'_h, \overline{P_h}) \in \mathbb{W}_h \times \mathbb{V}_h \times \mathbb{L}_h$, if one starts the procedure with a function $\phi'_{0,h}$, satisfying $\bar{\phi}'_{0,h} = 0$.*

Proof. Firstly, one proves immediately that the iterative sequence is well-posed, in particular, that the matrix in the linear system (2.23a) is invertible. For this, let $f = 0$ and $\bar{\phi}_h = 0$ in equation (2.21), moreover let $\psi'_h = \phi'_h$ and $\bar{Q}_h = \bar{P}_h$, we reduce $a(\phi'_h, \phi'_h) = 0$, and thus $\phi'_h = 0$ and $b(\bar{P}_h, \psi'_h) = 0$ for all $\psi'_h \in \mathbb{W}_h$. Finally we take $\psi'_h = \bar{P}_h$, which implies $\bar{P}_h = 0$. Thus the matrix of (2.23a) is invertible.
The convergence of this iterative sequence is immediate using the arguments of the continuous case. □

2.5.3 Comparison of the different AP-discretization matrices

In the equation (2.23b), A, denoted in the sequel by $\mathcal{M}_1$, is the matrix associated with the singular perturbation model, discretized by the finite element method introduced above. The matrix associated with the equation (2.23b), providing the fluctuating part and the Lagrangian approximation in the iterative method, will be denoted by $\mathcal{M}_2$. Finally we introduce $\mathcal{M}_3$ as the matrix associated with the solution of the whole system providing the discrete approximations $(\phi'_h, \bar{\phi}_h, \bar{P}_h)$. The structures and the sizes of these matrices are displayed on Figure 1. For completeness, the matrix associated with the original AP-scheme (2.9),(2.11), denoted by $\mathcal{M}_O$, is also included in this comparison.



(a) $\mathcal{M}_1 = A$ (b) $\mathcal{M}_2$ (c) $\mathcal{M}_3$ (d) $\mathcal{M}_O$

Mat.	$\mathcal{M}_1 = A$	$\mathcal{M}_2 = \begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix}$
Size	$N_x(N_z+2)$	$N_x(N_z+3)$
Nnz	$(3N_z+4)(3N_x-2)$	$(5N_z+8)(3N_x-2)$
Mat.	$\mathcal{M}_3 = \begin{pmatrix} A & \varepsilon C & B \\ \varepsilon C^T & \varepsilon A_2 & 0 \\ B^T & 0 & 0 \end{pmatrix}$	$\mathcal{M}_O = \begin{pmatrix} \tilde{A} & B \\ B^T & 0 \end{pmatrix}$
Size	$N_x(N_z+4)$	$N_x(N_z+3)$
Nnz	$(7N_z+13)(3N_x-2)$	$(N_z^2+6N_z+8)(3N_x-2)$

Figure 1: Structure (non-zero elements (Nnz)) and size of the discretization matrices ($\mathbb{Q}_1$ finite element method) for a grid size $(N_x, N_z) = (5, 5)$: (a) matrix associated with the singular perturbation problem (2.2), (b) matrix associated with the reformulated fluctuation equation(2.23b), (c) matrix associated with the direct resolution of the AP-scheme (2.22), (d) matrix associated with the fluctuation equation of the original AP formulation (2.11).

The matrix $\mathcal{M}_2$ is easily constructed thanks to the matrix of the singular perturbation problem by adding the two blocks related to the Lagrangian multiplier, corresponding to the zero-mean constraint. The discretization of this fluctuation equation is thus straightforward, requiring small localized modifications of the code providing the singular perturbation problem. The matrix $\mathcal{M}_O$ derived from the original AP-scheme has the same size as $\mathcal{M}_2$ but with a significantly larger number of non-zero coefficients. This is easily explained by the contribution of the integro-differential operator in the left hand side of the fluctuation equation in (2.11). In the new formulation this operator is moved to the second member providing a more sparse linear system. This underlines the essential advantages of the new AP-formulation introduced in this paper as compared to the one of [10]. Using the direct resolution increases the size of the linear system, since it provides the approximation for both components of the solution and for the Lagrangian. Note, however, that this increase of the matrix size is not dramatic, since both Lagrangian and mean part do not depend on z. The size of the blocks added in the matrix, namely B and C, is thus small compared with the size of A. Naturally both AP-formulation require the resolution of larger linear systems than the initial SP-problem. However, the crucial advantage is that no additional numerical effort is needed when $\varepsilon \to 0$, which is not the case for the SP-problem. In the latter case, the number of grid points has to increase with $\varepsilon \to 0$ in order to get the desired accuracy.

To highlight the arguments stated above with some concrete examples, we compared in Table 1, for two example cases $(N_x, N_z) = (50, 50)$ resp. $(N_x, N_z) = (500, 500)$, the four different matrices $\mathcal{M}_1$, $\mathcal{M}_2$, $\mathcal{M}_3$ and $\mathcal{M}_O$. What can be observed is that the $\mathcal{M}_O$-matrix corresponding to the AP-scheme of [10] is 11 times, even 101 times in the 500×500 case, more filled than the corresponding $\mathcal{M}_2$-matrix. This is rather drastic and permits to demonstrate the advantages of the here introduced AP-scheme as compared to the previous one. Interesting to observe is also that the ratio between matrix $\mathcal{M}_1$ and matrix $\mathcal{M}_2$ resp. between matrix $\mathcal{M}_1$ and matrix $\mathcal{M}_3$ is almost invariant, which means that the numerical efforts in solving the SP-model or the AP-formulation are rather the same.

		$\mathcal{M}_1$	$\mathcal{M}_O$	$\mathcal{M}_2$	$\mathcal{M}_3$
		SP-model	Original AP	Iterative AP	Direct AP
50×50	$p_1(\mathcal{M})$	0.34%	5.92%	0.54%	0.74%
	$p_2(\mathcal{M})$	1	18.2338	1.6753	2.3571
500×500	$p_1(\mathcal{M})$	0.00358%	0.6%	0.00594%	0.00829%
	$p_2(\mathcal{M})$	1	168.2234	1.6676	2.3358

Table 1: Comparison of the number of non-zero elements (Nnz) in the linear systems associated with the discrete Singular Perturbation problem, the original AP-scheme, and both the iterative and direct New AP-schemes, *i.e.* $p_1(\mathcal{M}) = \mathrm{Nnz}(\mathcal{M})/\mathrm{rk}(\mathcal{M})^2$, $p_2(\mathcal{M}) = \mathrm{Nnz}(\mathcal{M})/\mathrm{Nnz}(\mathcal{M}_1)$ where $\mathcal{M}$ is $\mathcal{M}_1$, $\mathcal{M}_O$, $\mathcal{M}_2$ and $\mathcal{M}_3$ respectively and $\mathrm{rk}(\mathcal{M})$ denotes the rank of matrix $\mathcal{M}$.

2.5.4 Numerical investigation of the new AP formulation

In this subsection the new AP-formulation (2.17) of the singular perturbation problem (2.2) is studied. In particular, we wish to demonstrate that the new formulation provides

the same properties as compared with the original one, introduced in [10], however with some major advantages. To this end, the numerical experiment proposed in [10] is reproduced. It consists in manufacturing an analytical set up for the problem, with an exact solution denoted ϕ_e and defined as

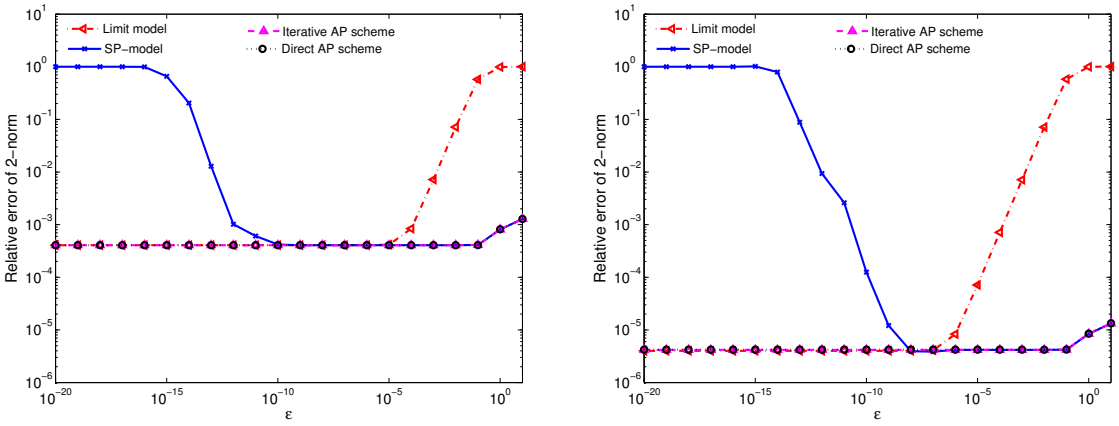
$$\phi_e(x,z) := \sin\left(\frac{2\pi}{L_x}x\right)\left(1+\varepsilon\cos\left(\frac{2\pi}{L_z}z\right)\right). \tag{2.24}$$

The coefficients of the elliptic problem (2.2) (verifying hypothesis 1) are defined as follows $A_\perp(x,z) = c_1 + xz^2$, $A_z(x,z) = c_2 + xz$, with two constants $c_1 > 0$, $c_2 > 0$. The right-hand side f of the problem is analytically computed by injecting the exact solution ϕ_e into (2.2). An approximation of this function ϕ_h can then be computed thanks to the different numerical methods introduced above, their precision being analyzed thanks to the relative error

$$\|\phi_e - \phi_h\|_r = \frac{\|\phi_e - \phi_h\|_2}{\|\phi_e\|_2}. \tag{2.25}$$

Note that $\|\phi_e\|_2^2 = \frac{1}{2}(1+\frac{\varepsilon^2}{2})LxLz$, thus this norm does not vanish when $\varepsilon \to 0$. For these experiments the simulation domain $[0,1]\times[0,1]$ is discretized by a uniform mesh with either 50×50 or 500×500 points, a $\mathbb{Q}_1$-finite element method, and a three-point Gauss-Legendre quadrature formula for the integral-discretizations.

On figure 2 the accuracy of the singular perturbation model (2.2), the New AP-formulation (2.9),(2.10) and the limit problem (2.5) are compared. For the AP-scheme both resolutions (the iterative (2.23) and the direct (2.22) ones) are considered.



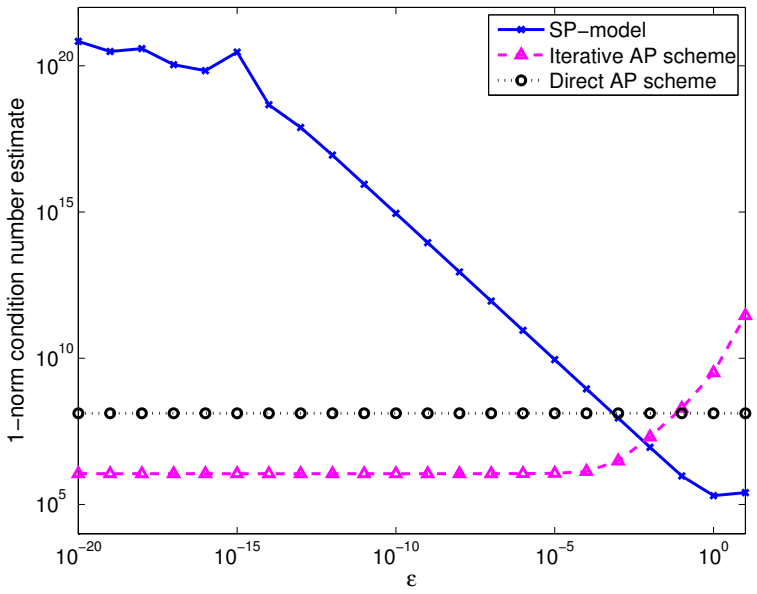
(a) (b)

Figure 2: Relative error between the exact solution and its approximations computed thanks to the limit model (2.5), the original (singular perturbation) model (2.2) as well as the iterative (2.23) and the direct new AP schemes (2.22) for meshes of (a): 50×50 and (b): 500×500 nodes.

One can observe that the limit model provides accurate approximations only for small values of the anisotropy ratio ε. For large enough ε-values the fluctuating part can not be neglected and a more complete model has to be used. The singular perturbation model

provides accurate results only for large ε-values, the precision of the approximation deteriorates significantly for ε smaller than 10^{-10} on the coarsest grid (see figure 2(a)). The domain of validity (accuracy) of the SP-scheme becomes even smaller on the refined mesh (see figure 2(b)), that is because the condition number of the refined mesh problem is larger than that of coarser grid problem. For example, when $\varepsilon = 10^{-12}$, the condition number for mesh 50×50 is 3.6387×10^{15}, while it is 3.3079×10^{17} for mesh 500×500. The accuracy of the solution, computed by the AP-scheme, is almost ε-independent demonstrating thus the efficiency of this new AP formulation for all anisotropy strength.

The condition number of the different linear systems is plotted on figure 3(a), as a function of ε. It is computed thanks to the block algorithm for matrix 1-norm estimation [16]. The new AP-scheme provides the same advantageous features as the original



(a)

ε		1	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-8}	10^{-9}	10^{-20}
Dire.	r_D	1.8	1.8	1.8	1.8	1.8	1.8	1.8	1.8	1.8	1.8	1.8
Iter.	r_I	6.0	5.4	5.0	4.0	3.4	2.2	2.0	2.0	1.7	1.7	1.7
	n_I	18	16	14	11	8	4	3	3	2	2	2

(b)

Figure 3: Comparison between the 2D original model (2.2), the iterative AP scheme (2.23) and the direct AP scheme (2.22) with mesh size 250×250. (a) Condition number estimate for the discretization matrices; (b) Computational time of the direct AP- resp. iterative AP-resolutions, divided by the computational time of the SP-model resolution, denoted by r_D, r_I respectively. The iteration number of iterative resolution, denoted by n_I, is also quoted. The linear systems are solved by the direct sparse solver PARDISO.

scheme, with a matrix whose condition number is almost independent of the ε-values for large anisotropy ratios. The discretized singular perturbation problem gives a matrix

whose condition number blows up with vanishing ε. This explains the poor accuracy of this scheme for large anisotropy ratios.

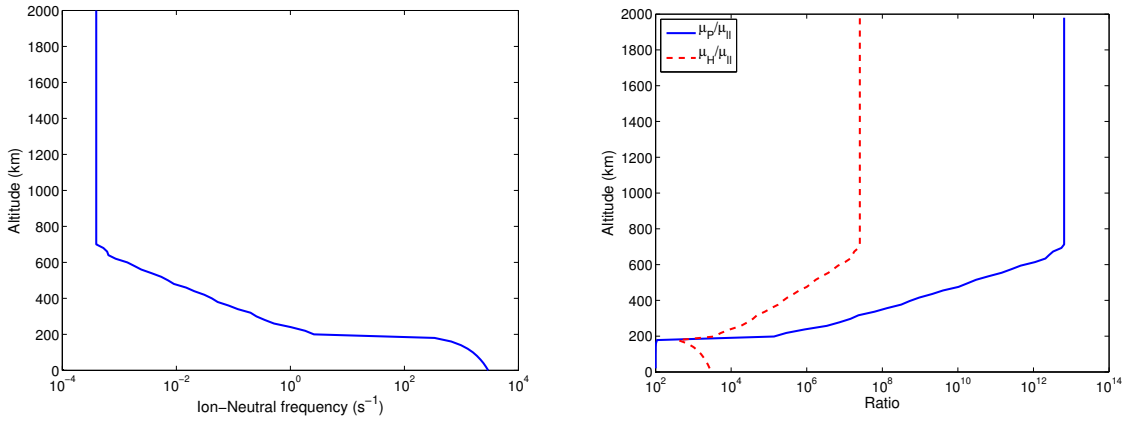
The computational efficiency, as a function of the anisotropy ratio, is evaluated in table 3(b) for the different approaches. The time required for the computation via the singular perturbation model is used as a reference. This computations are performed thanks to a direct sparse linear solver PARDISO [28], [29] and do not depend on the ε-values. For the direct resolution, the linear system is solved only one time. While for the iterative resolution, two (different size) linear systems are solved several times until the difference between two successive solutions of system (2.23) is small. Note that the most expensive part of the direct sparse linear solver is the matrix factorization. Thus in the iterative procedure we can store the factorized matrices, and use then these matrices in each iteration. Therefore we save a lot of computational-consuming.

A larger number of non zeros elements as well as an increased system size explains the larger amount of computations required by the AP-scheme. The direct resolution is roughly 80% slower than that of the SP-model. However, since the AP-reformulations provide matrices with a better conditioning, we expect the AP-scheme to be more efficient than the SP-model when a Krylov method is used as linear system solvers. This point is out of the scope of the present paper and is deferred to future work. Moreover, we note the efficiency of the iterative resolution. For the smallest ε-values its efficiency is comparable to that of the direct resolution. For the largest values, this approach is revealed to be computationally more demanding, due to a large number of iterations required to reach the convergence. For $\varepsilon = 1$, 18 iterations are indeed necessary to compute an accurate approximation, giving rise to a computational effort almost three times as large as that of the direct resolution.

3 Towards a model problem well suited for plasma applications: heterogeneous anisotropy ratios

3.1 Motivation and design of a more general model problem

The second aim of this paper is to provide a numerical method able to handle anisotropic models arising in plasma physics problems and more specifically in ionospheric plasma simulations where the anisotropy of the medium is related to the earth magnetic field. In this context, the ion-neutral collisions are responsible for a large particle mobility [5], [12] along the magnetic field lines, whereas the transverse one is rather small. However this collision frequency undergoes huge variations with the altitude, as depicted on figure 4(a), which can be as huge as ten orders of magnitude on an altitude range of one thousand kilometers ([7], [18]). The anisotropy of this parameter in the ionosphere is thus very large at high altitudes while vanishing for the lowest ones (see figure 4(b)).



(a) (b)

Figure 4: Typical ionospheric plasma main characteristics as a function of the altitude. (a) Ion-neutral collision frequency. (b) Ratio of the transverse and aligned (with respect to the earth magnetic field) particle mobility. μ_P, μ_H, $\mu_{\|}$ denote the Pedersen, Hall, field-aligned mobilities respectively [2].

The model problem considered so far is representative for the ionosphere in the highest altitudes. In order to handle this huge variation of parameters in the whole range of altitudes, the model considered in the sequel will use a non homogeneous anisotropy ratio, depending on the space coordinates (x, z). This refined model problem consists in finding $\phi(x, z)$ verifying

$$\begin{cases} -\dfrac{\partial}{\partial x}\left(A_\perp(x,z)\dfrac{\partial\phi(x,z)}{\partial x}\right) - \dfrac{\partial}{\partial z}\left(\dfrac{A_z(x,z)}{\varepsilon(x,z)}\dfrac{\partial\phi(x,z)}{\partial z}\right) = f(x,z), \quad \text{on } \Omega, \\ \phi(x,z) = 0, \qquad (x,z) \in \partial\Omega_x \times \Omega_z, \\ \partial_z\phi(x,z) = 0, \qquad (x,z) \in \Omega_x \times \partial\Omega_z, \end{cases} \tag{3.1}$$

where $A_\perp$, A_z and ε are positive functions, the latter presenting steep gradients. The system (3.1) is a simplified version of the so called Dynamo-3D model [4], [22] providing the electric potential in the quasi-neutral ionospheric plasma.

3.2 Asymptotic preserving formulation for heterogeneous anisotropy ratios

In this subsection, the asymptotic-preserving reformulation is derived for non homogeneous anisotropy ratios. Keeping in mind that ε is a function of the space variables, the formulations of the average and fluctuation equations is now written as (for comparison see (2.9),(2.10))

$$\begin{cases} -\frac{\partial}{\partial x}\left(\overline{A_\perp}\frac{\partial\bar\phi}{\partial x}\right) = \bar f + \frac{\partial}{\partial x}\left(\overline{A_\perp\frac{\partial\phi'}{\partial x}}\right), & \text{in } \Omega_x, \\ \bar\phi = 0, & \text{on } \partial\Omega_x. \end{cases} \tag{3.2a}$$

$$\begin{cases} -\frac{\partial}{\partial x}\left(A_\perp \frac{\partial \phi'}{\partial x}\right) - \frac{\partial}{\partial z}\left(\frac{A_z}{\varepsilon}\frac{\partial \phi'}{\partial z}\right) = f + \frac{\partial}{\partial x}\left(A_\perp \frac{\partial \bar{\phi}}{\partial x}\right), & \text{in } \Omega, \\ \phi' = 0, & \text{on } \partial\Omega_x \times \Omega_z, \\ \partial_z \phi' = 0, & \text{on } \Omega_x \times \partial\Omega_z, \\ \overline{\phi'} = 0, & \text{in } \Omega_x. \end{cases} \tag{3.2b}$$

The average equation (3.2a) is the one derived for the homogeneous ε-case, the fluctuation equation (3.2b) being slightly modified. We introduce the same Hilbert-spaces $\mathbb{V}$, $\mathbb{W}$, as defined in section 2.2, and a new ε-independent scalar product

$$(\phi, \psi)_{\mathbb{V}} = (\partial_x \phi, \partial_x \psi)_{L^2(\Omega)} + (\partial_z \phi, \partial_z \psi)_{L^2(\Omega)}.$$

By redefining some terms of the bilinear forms (2.13), especially $a_0(\cdot,\cdot)$, $a(\cdot,\cdot)$ and $b(\cdot,\cdot)$

$$\begin{aligned} a_0(\phi', \psi') &:= \int_0^{L_z}\int_0^{L_x} \frac{A_z(x,z)}{\varepsilon(x,z)} \frac{\partial \phi'}{\partial z}(x,z)\frac{\partial \psi'}{\partial z}(x,z)dxdz, \\ a_1(\phi', \psi') &:= \int_0^{L_z}\int_0^{L_x} A_\perp(x,z) \frac{\partial \phi'}{\partial x}(x,z)\frac{\partial \psi'}{\partial x}(x,z)dxdz, \\ a_2(\bar{\phi}, \bar{\psi}) &:= \int_0^{L_x} \bar{A}_\perp(x) \frac{\partial \bar{\phi}}{\partial x}(x)\frac{\partial \bar{\psi}}{\partial x}(x)dx, \\ a(\phi', \psi') &:= a_0(\phi', \psi') + a_1(\phi', \psi'), \\ b_1(\bar{P}, \psi') &:= \int_0^{L_x} \bar{P}(x) \int_0^{L_z} \frac{1}{\varepsilon(x,z)} \psi'(x,z)dzdx, \\ b_2(\phi', \bar{Q}) &:= \frac{1}{L_z}\int_0^{L_x} \bar{Q}(x) \int_0^{L_z} \phi'(x,z)dzdx, \\ c(\bar{\phi}, \psi') &:= \int_0^{L_z}\int_0^{L_x} A_\perp(x,z) \frac{\partial \bar{\phi}}{\partial x}(x)\frac{\partial \psi'}{\partial x}(x,z)dxdz, \end{aligned} \tag{3.3}$$

the weak formulation of (3.2) writes now

$$\begin{cases} a_2(\bar{\phi}, \bar{\psi}) = (\bar{f}, \bar{\psi}) - \frac{1}{L_z} c(\bar{\psi}, \phi'), & \forall \bar{\psi} \in \mathbb{W}, \\ a(\phi', \psi') = (f, \psi') - c(\bar{\phi}, \psi'), & \forall \psi' \in \mathbb{V}. \end{cases} \tag{3.4}$$

The AP-reformulation of problem (3.4) is again deduced by introducing a Lagrange multiplier $\bar{P}$, corresponding to the constraint $\bar{\phi}' \equiv 0$. We thus have

$$(\text{AP})_{var} \begin{cases} a_2(\bar{\phi}, \bar{\psi}) = (\bar{f}, \bar{\psi}) - \frac{1}{L_z} c(\bar{\psi}, \phi'), & \forall \bar{\psi} \in \mathbb{W}, \\ a(\phi', \psi') + b_1(\bar{P}, \psi') = (f, \psi') - c(\bar{\phi}, \psi'), & \forall \psi' \in \mathbb{V}, \\ b_2(\bar{Q}, \phi') = 0, & \forall \bar{Q} \in L^2(\Omega_x). \end{cases} \tag{3.5}$$

Remark 3.1. *Note that the discrete linear system obtained for the fluctuation-part is not symmetric for non homogeneous anisotropy ratios. The rescaling of the fluctuation equation (by ε), introduced for constant ε, cannot be reproduced anymore. Accordingly, the bilinear form b defined in (2.13) by*

$$b(\bar{P}, \psi') := \int_0^{L_x} \bar{P}(x) \int_0^{L_z} \psi'(x,z)dzdx$$

is now recasted into

$$b_1(\bar{P},\psi') := \int_0^{L_x} \bar{P}(x) \int_0^{L_z} \frac{1}{\varepsilon(x,z)} \psi'(x,z) dz dx.$$

With this new definition, all the terms appearing in the second equation of (3.5) scale as $1/\varepsilon$ *in the limit* $\varepsilon \to 0$*, which guaranties the asymptotic preserving property of the scheme. Note that if we assume a homogeneous* ε*, the fluctuation equation in* (3.5) *can be rescaled by* ε *and the system recovers its previous symmetric structure* (2.13).

We shall now prove that the system (3.5) is equivalent to the problem (3.4).

Hypothesis 3. *Let in the following* $\varepsilon \in L^\infty(\Omega)$*, satisfying* $0 < \varepsilon_0 \leq \varepsilon(x,z) \leq \varepsilon_M \leq 1$ *with* ε_0 *resp.* ε_M *two constants.*

Proposition 3.2. *Under hypothesis 1 and hypothesis 3, there exists a unique* $(\bar{\phi},\phi',\bar{P}) \in \mathbb{W}\times\mathbb{V}\times L^2(\Omega_x)$ *satisfying (3.5). Moreover, the pair* $(\bar{\phi},\phi')$ *is the unique solution of (3.4) and* $\bar{P}\equiv 0$.

Proof. The well-posedness of system (3.4) can be proved as we did in the constant ε-case. So it remains to prove the equivalence between the system (3.4) and system (3.5). First let $(\bar{\phi},\phi') \in \mathbb{W}\times\mathbb{V}$ be the unique solution of system (3.4), then $(\bar{\phi},\phi',0)$ will verify system (3.5). Inversely, we assume $(\bar{\phi},\phi',\bar{P}) \in \mathbb{W}\times\mathbb{V}\times L^2(\Omega_x)$ to be a solution of system (3.5). Taking in the second equation of system (3.5), test functions ψ' depending only on x, leads to

$$\begin{aligned} & L_z \int_0^{L_x} \overline{A_\perp \partial_x \phi'} \partial_x \psi' dx + \int_0^{L_x} \bar{P}(x)\psi'(x) \int_0^{L_z} \frac{1}{\varepsilon(x,z)} dz\, dx \\ = \; & L_z \int_0^{L_x} \bar{f}\psi' dx - L_z \int_0^{L_x} \overline{A}_\perp \partial_x \bar{\phi} \partial_x \psi' dx. \end{aligned}$$

It is easy to see that $L_z \int_0^{L_x} \overline{A_\perp \partial_x \phi'} \partial_x \psi' dx + L_z \int_0^{L_x} \overline{A}_\perp \partial_x \bar{\phi} \partial_x \psi' dx = L_z \int_0^{L_x} \bar{f}\psi' dx$ according to the first equation of system (3.5). Thus we get

$$\int_0^{L_x} \bar{P}(x)\psi'(x) \int_0^{L_z} \frac{1}{\varepsilon(x,z)} dz\, dx = 0, \;\; \forall \psi' \in \mathbb{W}\,.$$

As $\int_0^{L_z} \frac{1}{\varepsilon(x,z)} dz > 0$, we deduce that $\bar{P} \equiv 0$, by density arguments. □

3.3 Steep gradients of the heterogeneous anisotropy ratio

Before discretizing the AP-system (3.5), we first have to think about the numerical problems arising from rapid variations of the function $\frac{1}{\varepsilon}$. As we know, a rapidly varying continuous function can be considered numerically as a discontinuous function in certain intervals of a grid. One can refine the discretization mesh to obtain a preciser approximate solution, but with more computational efforts. For example, in 3-dimensional

simulations, the computation complexity is equal to N^3, where N is the number of intervals in each direction. A refinement may not be acceptable for large N.

To overcome this difficulty of discretizing abrupt gradients, we shall be guided by the ideas proposed by Saito *et al.* [31]. It considers a two-point boundary value problem of the form

$$-(pu')' = f(x \in I)\,, \quad u|_{x=0,1} = 0,$$

where $p \in L^\infty$ and $f \in L^2$ are given functions. The gradient of p is very steep. Then Saito *et al.* [31] studies the error estimate of a standard finite element method via the Green function of the solution. In fact, denoting by $AU = F$ the linear system corresponding to the finite element method, we note that $G := A^{-1}$ is an approximation of this Green function. Thus Saito *et al.* [31] proposed to use a harmonic mean to approximate the Green function. In more details, the coefficients of the FE-discretization matrix A are given by $a_i = \dfrac{1}{\Delta x^2}\displaystyle\int_{x_{i-1}}^{x_i} p\,dx$, while their harmonic mean is $\tilde{a}_i = \dfrac{1}{\Delta x^2}\left(\displaystyle\int_{x_{i-1}}^{x_i} p^{-1}dx\right)^{-1}$. The error analysis was presented by T. Tsuchiya *et al.* [34]. It shows that the harmonic mean method is more precise than the arithmetic mean in the case where p is not so regular, *i.e.* when the gradient of p is very large. Moreover, the Scharfetter-Gummel (SG) scheme, which is a special quadrature-formula of the harmonic mean, is shown to have an error estimate depending only on $\ln p$ [31]. Thus the SG scheme is expected to give more precise numerical results as compared with the standard finite element method especially when p is not so regular.

Let us thus propose here three approaches to handle with the steep gradient in problem (3.5). We shall consider, for sake of simplicity, anisotropies $\varepsilon(z)$ depending only on z.
The first approach is to use a non-conservative formulation for the fluctuation equation. By developing $\frac{\partial}{\partial z}\left(\frac{A_z}{\varepsilon}\frac{\partial \phi'}{\partial z}\right)$ in (3.2b), we obtain the following equation

$$-\varepsilon\frac{\partial}{\partial x}\left(A_\perp\frac{\partial\phi'}{\partial x}\right) - \frac{\partial}{\partial z}\left(A_z\frac{\partial\phi'}{\partial z}\right) + \frac{\partial(\ln\varepsilon)}{\partial z}A_z\frac{\partial\phi'}{\partial z} = \varepsilon f + \varepsilon\frac{\partial}{\partial x}\left(A_\perp\frac{\partial\bar{\phi}}{\partial x}\right).$$

To implement this approach, we use again the bilinear forms (2.13), but change only $a(\phi', \psi')$ in

$$a_0(\phi', \psi') := \int_0^{L_z}\int_0^{L_x} A_z(x,z)\frac{\partial\phi'}{\partial z}(x,z)\frac{\partial\psi'}{\partial z}(x,z) + \frac{\partial \ln\varepsilon(z)}{\partial z}A_z(x,z)\frac{\partial\phi'}{\partial z}(x,z)\psi'(x,z)dxdz.$$

However, this expression is no more in a conservative form.

Secondly, we use the harmonic mean method. That is to consider $\frac{A_z}{\varepsilon}$ in (3.5) as the rapidly varying function p, and we use the harmonic mean form $\tilde{a}$ proposed above.
The third approach is the Scharfetter-Gummel scheme proposed in [31]. We refer the reader to appendix A for more details.
In the sequel we shall use this three approaches to discretize problem (3.5) and compare them with a standard FE-discretization of (3.5).

3.4 Numerical results

To verify the efficiency of the AP-scheme for highly heterogeneous anisotropy ratio problems, we choose an artificially constructed variable ε test case. Obviously, it does not correspond to the physical ratios. However, if the AP-scheme behaves well for these formal test cases, it is also expected to be efficient for the real physical problems. We use again the exact solution and the diffusion matrix of the constant ε case, but replace the constant ε by a variable on of the form

$$\varepsilon_1(z) = \begin{cases} \frac{1}{2}\left(\varepsilon_{\max}\left(1+\tanh\left(q\left(0.1L_z - z\right)\right)\right) + \varepsilon_{\min}\left(1-\tanh\left(q\left(0.1L_z - z\right)\right)\right)\right), \\ \qquad\qquad \text{if } 0 \le z \le \frac{L_z}{2}, \\ \frac{1}{2}\left(\varepsilon_{\max}\left(1+\tanh\left(q\left(z - 0.9L_z\right)\right)\right) + \varepsilon_{\min}\left(1-\tanh\left(q\left(z - 0.9L_z\right)\right)\right)\right), \\ \qquad\qquad \text{if } \frac{L_z}{2} \le z \le L_z. \end{cases} \tag{3.6}$$

The variable ε_1 is controlled by three parameters q, $\varepsilon_{\max}$, $\varepsilon_{\min}$: q describes the steep slope of the curve, $\varepsilon_{\max}$, $\varepsilon_{\min}$ control the maximum and minimum values of ε_1. An example of $\varepsilon_1(z)$ is illustrated in figure 5, it is constructed so, to sketch the variation of the physical parameters presented on figure 4(b).

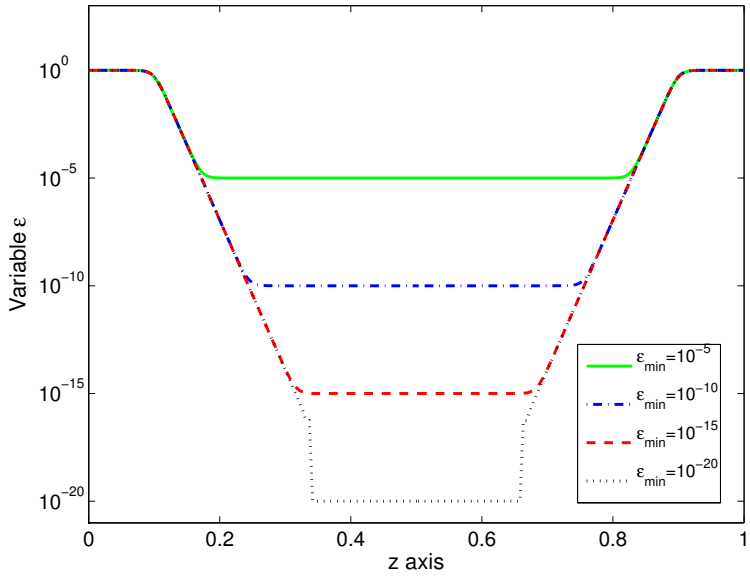


Figure 5: Variable ε_1 for different $\varepsilon_{\min}$. We fix $q = 80$, $\varepsilon_{\max} = 1$ and draw the curve of ε_1 for several $\varepsilon_{\min}$ equal to 10^{-5}, 10^{-10}, 10^{-15}, 10^{-20} respectively.

In the following numerical tests, we fix the parameters $\varepsilon_{\max} = 1$, $q = 80$ and vary the parameter $\varepsilon_{\min}$ from 10^{-20} to 1. The right-hand side f is obtained by injecting $\phi_e(x, z)$ into equation (3.1). We shall compare the numerical results obtained by all the methods mentioned above: 2D SP-model (3.1) discretized by the $\mathbb{Q}_1$ finite element method, the standard AP scheme (3.5), the non-conservative AP scheme, the harmonic mean AP scheme, and the Scharfetter-Gummel AP scheme. The four AP schemes are solved by

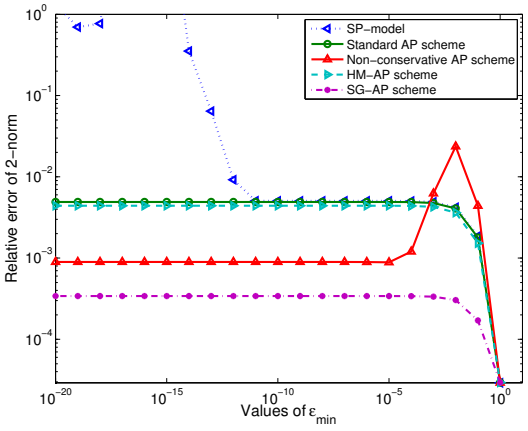
both iterative and direct resolutions, however the relative errors of these two resolutions are the same. Thus we only present the numerical results of the direct resolution.

In figure 6(a), we observe that the relative error in the L^2-norm of the 2D SP-model increases with vanishing $\varepsilon_{\min}$ values. This error remains close to 10^{-3} for $\varepsilon_{\min}$ between 10^{-3} and 10^{-11}, but explodes for $\varepsilon_{\min} \leq 10^{-11}$. The curve of the standard AP-scheme coincides with that of the SP-model when $\varepsilon_{\min} > 10^{-11}$, the accuracy of the former on remaining $\varepsilon_{\min}$ independent for larger anisotropy ratios. The non-conservative AP scheme demonstrates a non monotone accuracy evolution, with a peak obtained for $\varepsilon_{\min} = 10^{-2}$. However, the error norm is close to 2×10^{-4} for $\varepsilon_{\min}$ below 10^{-5}. The curve of the harmonic mean AP scheme is similar to that of the standard AP scheme but with smaller relative error. Finally, the Scharfetter-Gummel AP scheme gives the best accuracy with relative error one order of magnitude smaller than that of the standard AP scheme.

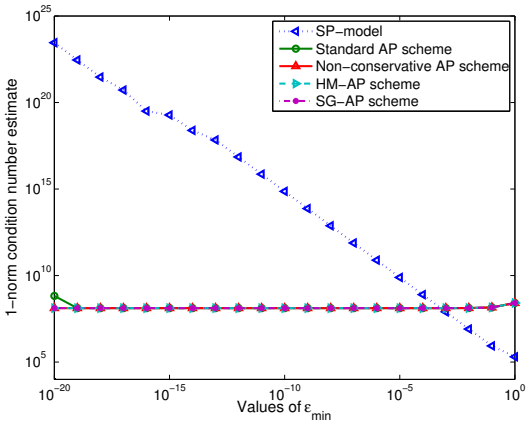
On the figure 6(b) the condition number estimates of the different methods are displayed. These conditioning estimations are computed after an equilibration of the matrices is performed. This procedure consists in multiplying $\mathcal{M}_2$, respectively $\mathcal{M}_3$, by the row balance matrix $\mathcal{P}_2$, respectively $\mathcal{P}_3$, defined as

$$\mathcal{P}_2 = \begin{pmatrix} \mathcal{E}_0 & & & \\ & \ddots & & \\ & & \mathcal{E}_{N_z+1} & \\ & & & I \end{pmatrix}, \qquad \mathcal{P}_3 = \begin{pmatrix} \mathcal{E}_0 & & & \\ & \ddots & & \\ & & \mathcal{E}_{N_z+1} & \\ & & & I \\ & & & & I \end{pmatrix},$$

where $\mathcal{E}_k = \varepsilon(z_k)I$ for $0 \leq k \leq N_z + 1$ and I is the $N_x \times N_x$ identity matrix. As in the homogeneous ε case, the condition number of the SP-model increases with the anisotropy ratio, while the curves corresponding to the four AP schemes almost coincide and remain quite independent of the $\varepsilon_{\min}$ values.



(a) Relative error between the exact solution and its approximations, as a function of $\varepsilon_{\min}$



(b) Condition number estimate of the equilibrated matrices, as a function of $\varepsilon_{\min}$

Figure 6: Comparison of the SP-model, the standard AP scheme, the non-conservative AP scheme, the harmonic mean AP scheme and the Scharfetter-Gummel AP scheme on a 2 Dimensional computation carried out on a 250×250 mesh, using a $\mathbb{Q}_1$ finite element method.

Next we investigate the efficiency of the iterative and direct resolution of AP-scheme. In table 2 both resolutions are compared for computations carried out on meshes with 50×50, 250×250 and 500×500 cells. The anisotropic ratio is computed as before with $\varepsilon_{\max} = 1$, $q = 80$ and $\varepsilon_{\min}$ values ranging from 10^{-20} to 1. For the iterative resolution, the sequence is initiated with $\phi' = x(x - L_x)\cos\left(\frac{2\pi z}{L_z}\right)$. We note that the number of iterations required to reach the convergence increases weakly with the mesh size, but significantly with the values of the anisotropy ratio. The number of iterations required for the smallest values of $\varepsilon_{\min}$ is three to four times less than that of the isotropy configuration. For the two dimensional tests investigated so far, the direct resolution, although producing a larger linear system, demonstrates to be more efficient than the iterative one. The table 2 also displays the computational times for both approaches related to that of the SP-model. The efficiency of the direct approach does not deteriorate significantly with the mesh refinement. Indeed with a mesh composed of 50×50 cells, the AP scheme is 40% slower than the SP-model. Using a more refined mesh, with one hundred times the number of cells, the AP scheme is roughly two times slower than the SP-model. The iterative resolution requires more computational time, between 2 and 4 times, than the direct one. Note however that the relative efficiency of this two resolutions may be altered if a good estimate of the solution can be used to initiate the sequence.

$\varepsilon_{\min}$			1	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-10}	10^{-20}
50×50	Iter.	n_I	17	13	10	7	5	5	5	5
		r_I	6.5	5.3	4.6	4.3	3.5	3.5	3.5	3.5
	Dire.	r_D	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4
250×250	Iter.	n_I	18	15	12	10	7	5	5	5
		r_I	7.7	6.8	6.1	5.5	4.8	4.2	4.2	4.2
	Dire.	r_D	1.8	1.8	1.8	1.8	1.8	1.8	1.8	1.8
500×500	Iter.	n_I	19	15	13	11	8	6	5	5
		r_I	7.3	6.4	5.9	5.4	4.8	4.4	4.1	4.1
	Dire.	r_D	2.1	2.1	2.1	2.1	2.1	2.1	2.1	2.1

Table 2: Computational efficiency of the the iterative and direct resolutions of the AP-scheme on meshes with 50×50, 250×250 and 500×500 cells: n_I is the number of iteration of iterative resolution, r_I (resp. r_D) is the computional time of iterative (resp. direct) resolution divided by the computational time of the SP-model. All the linear system are solved thanks to a sparse linear direct solver [28].

4 A 3-D physical test case

4.1 Introduction

The aim of this section is to generalize the just introduced AP-scheme in order to apply it to a real 3D ionospheric plasma problem [4]. In fact, in this model the diffusion matrix $\mathcal{A}$ is of the following form

$$\mathcal{A} = \begin{pmatrix} \mu^P & -\mu^H & 0 \\ \mu^H & \mu^P & 0 \\ 0 & 0 & \mu^{\parallel} \end{pmatrix}, \tag{4.1}$$

where μ^P, μ^H, $\mu^{\parallel}$ are the Pedersen, Hall and field-aligned mobilites respectively [2]. Thanks to standard ionospheric models (see for instance the IRI model [7] or the SAMI2 model [18], [19]) these quantities can be estimated, demonstrating large differences of μ^P and μ^H magnitudes as compared to that of $\mu^{\parallel}$, shown in figure 4(b). Moreover, the ratio between μ^P/μ^H as a function of the altitude can also be very large in certain ionospheric layer. In another words, the anisotropy variations are not limited to one directions. In this section, we will thus focus on a 3-dimensional anisotropic elliptic problem and demonstrate that the AP scheme is valid for more complicated and realistic heterogeneous anisotropic problem.

4.2 3-dimensional model

Let us consider the following 3-dimensional anisotropic elliptic problem in $\Omega \subset \mathbb{R}^3$

$$\begin{cases} -\nabla \cdot (\mathcal{A}\nabla\phi) = f, \text{ in } \Omega, \\ \mathcal{A}\nabla\phi \cdot \overrightarrow{n} = 0, \text{ on } \partial\Omega_x \times \Omega_y \times \Omega_z \cup \Omega_x \times \Omega_y \times \partial\Omega_z, \\ \phi = 0, \text{ on } \Omega_x \times \partial\Omega_y \times \Omega_z, \end{cases} \tag{4.2}$$

where $\mathcal{A}$ is the diffusion matrix of the form

$$\mathcal{A} = \begin{pmatrix} A & -\varepsilon D & 0 \\ \varepsilon D & B & 0 \\ 0 & 0 & \frac{1}{\varepsilon} C \end{pmatrix} \tag{4.3}$$

and $A(x,y,z)$, $B(x,y,z)$, $C(x,y,z)$ resp. $D(x,y,z)$ are known functions of the same order of magnitude. The parameter $0 < \varepsilon \leq 1$ can be a constant or a function of all variables provoking the anisotropy of the problem.

Let us first study the properties of problem (4.2)-(4.3). For this we define the following Hilbert space

$$\mathbb{V} = \{\phi \in H^1(\Omega) \ / \ \phi = 0 \text{ on } \Omega_x \times \partial\Omega_y \times \Omega_z\},$$

with corresponding scalar product

$$(\phi, \psi)_{\mathbb{V}} = (\partial_x \phi, \partial_x \psi)_{L^2(\Omega)} + (\partial_y \phi, \partial_y \psi)_{L^2(\Omega)} + (\partial_z \phi, \partial_z \psi)_{L^2(\Omega)},$$

and then write the weak formulation of (4.2) under the form

$$a(\phi, \psi) = (f, \psi), \quad \forall \psi \in \mathbb{V} \tag{4.4}$$

with $a(\phi, \psi) := \displaystyle\int_\Omega (\mathcal{A}\nabla\phi) \cdot \nabla\psi \, dxdydz$. We prove in the next proposition that (4.4) admits a unique weak solution under the following hypothesis.

Hypothesis 4. *Let the diffusion functions $A, B, C, D \in L^\infty(\Omega)$ as well as the anisotropy $\varepsilon \in L^\infty(\Omega)$ satisfy $0 < c_{\min} \leq A, B, C \leq c_{\max}$ resp. $0 < \varepsilon_0 \leq \varepsilon \leq \varepsilon_M \leq 1$ where $c_{\min}$, $c_{\max}$, ε_0, ε_M are positive constants. Moreover let $f \in L^2(\Omega)$.*

Proposition 4.1. *Under hypothesis hypothesis 4, the equation (4.4) has a unique weak solution $\phi \in \mathbb{V}$.*

Proof. To prove this proposition, we use the lemma of Lax-Milgram. Indeed, the coercivity of the bilinear form $a(\cdot,\cdot)$ is immediate, as

$$\begin{aligned} a(\phi, \phi) &= \int_\Omega A|\partial_x \phi|^2 + B|\partial_y \phi|^2 + \frac{1}{\varepsilon} C|\partial_z \phi|^2 dxdydz \\ &\geq c_{\min} \|\phi\|_{\mathbb{V}}^2, \ \phi \in \mathbb{V}. \end{aligned}$$

The continuity of $a(\cdot,\cdot)$ can be easily verified. □

The AP reformulation of equation (4.2) is just a generalization of (3.5). We leave this computation for the interested reader.

4.3 Numerical experiments

The procedure used so far is reproduced for these numerical investigations of the AP-schemes. The exact solution of the system is defined by

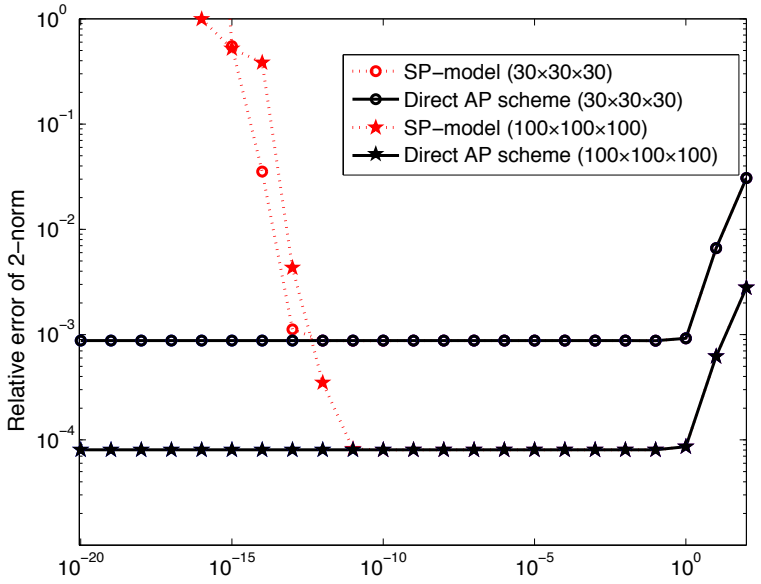
$$\phi_e(x,y,z) = x^2(L_x - x)^2 y^2 (L_y - y)^2 \left(1 + \varepsilon \sin^2\left(\frac{2\pi z}{L_z}\right) \cos^2\left(\frac{2\pi z}{L_z}\right)\right),$$

which satisfies the boundary conditions (4.2). The diffusion matrix coefficients are defined

$$A = c_1 + x^2yz\,, \quad B = c_2 + xy^2z\,, \quad C = c_3 + xyz^2\,, \quad D = c_4 + xyz\,.$$

The right-hand side f is computed analytically, by injecting ϕ_e into (4.2). This gives the setup of the different experiments studied in the sequel. The discretization techniques are readily extended from the 2-dimensional case detailed in subsection 3.3 using standard $\mathbb{Q}_1$ finite element methods.

Firstly, the constant ε case is investigated with the procedure used for the 2D case: the values of ε are varied from 10^{-20} to 10^2 and the error norm between the exact solution and the numerical approximations computed thanks to the SP-model and the standard AP-scheme are compared. These results are gathered in figure 7(a) for calculations carried out on mesh sizes $30 \times 30 \times 30$ and $100 \times 100 \times 100$. The SP-model is observed to fail in providing accurate approximations for $\varepsilon < 10^{-12}$ on the more refined mesh. The computational efficiency of the AP-schemes iterative and direct resolutions is also investigated in table 7(b). The conclusion of the 2D experiments still holds for the 3-Dimensional case, with a weak dependence of the iteration number accordingly to the mesh size, except for the largest values of ε. For this value a low convergence rate is observed. However, even in the more unfavorable configuration, the iterative resolution is the most efficient. This may be explained by the direct solver loss of efficiency for three dimensional elliptic problem with a dramatic filled-in of the factorized matrix [37], [35]. The iterative resolution allows to reduce the size of the linear systems which explain the relative efficiency of this approach compared to the direct resolution.



(a) Relative error between exact solution and its approximations

ε			10^2	10^1	10^0	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-10}	10^{-20}
$30 \times 30 \times 30$	Dire.	r_D	1.8	1.8	1.8	1.8	1.8	1.8	1.8	1.8	1.8
	Iter.	r_I	2.1	1.7	1.6	1.6	1.6	1.6	1.4	1.4	1.4
		n_I	7	4	3	3	3	3	2	2	2
$100 \times 100 \times 100$	Dire.	r_D	2.6	2.6	2.6	2.6	2.6	2.6	2.6	2.6	2.6
	Iter.	r_I	2.4	2.1	2.1	2.1	2.1	2.1	2.1	2.1	2.1
		n_I	22	3	2	2	2	2	2	2	2

(b) Ratio of computational time between resolutions of AP scheme and SP-model

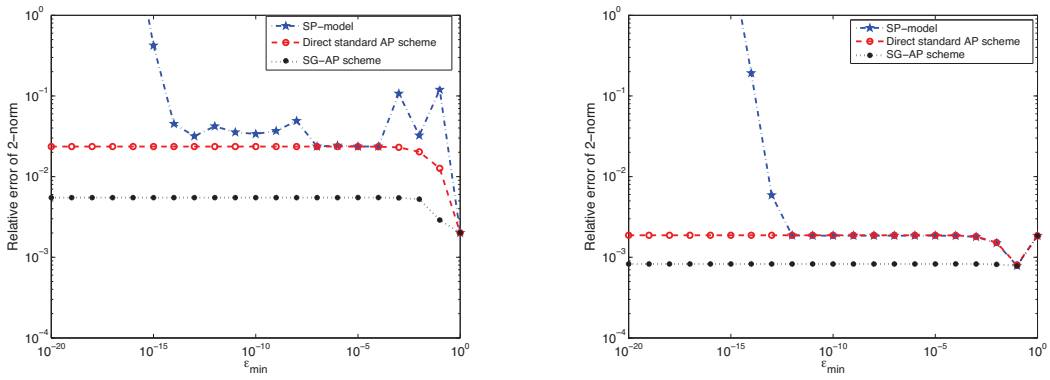
Figure 7: Comparison among the 3-dimensional SP-model and the AP scheme in the constant ε case (for both iterative and direct resolutions). The values of ε varies between 10^{-20} and 10^2. Two different sizes $30 \times 30 \times 30$ and $100 \times 100 \times 100$ are compared respectively. We denote r_I, r_D the ratio of computational time between iterative resp. direct AP scheme and SP-model, n_I the iteration number of iterative resolution.

Next, we consider a variable ε case which is representative of the physic case, *i.e.* the anisotropy variations are in all the directions and it is more notable in direction z. Such a variable ε is given as follows

$$\varepsilon_2(x, y, z) = \varepsilon_1(z) \frac{(x - x_{mid})^2 + (y - y_{mid})^2 + (z - z_{mid})^2 + 1}{x_{mid}^2 + y_{mid}^2 + z_{mid}^2 + 1}, \tag{4.5}$$

where $\varepsilon_1(z)$ is the same as (3.6), x_{mid}, y_{mid}, z_{mid} denote the middle point of interval in each direction. Similarly the parameters are set to $q = 80$, $\varepsilon_{\max} = 1$ and with $\varepsilon_{\min}$ values ranging from 10^{-20} to 1. The approximation accuracy of the SP-model, standard AP-scheme and SG-AP scheme is analyzed in figure 8(a) and 8(b). All the schemes provide the similar approximation qualities as for the 2D investigations: the SP-model being precise only for the largest $\varepsilon_{\min}$ values and the AP-schemes accuracy independent

of the anisotropy ratio. We note the iterative resolution is less precise than the direct resolution in this test case. The table 8(c) states that the computational efficiency of the AP-scheme, direct and iterative resolutions. These results confirm that, contrary to the 2D experiments, for the 3D case the direct resolution is more computational-consuming than the iterative one.



(a) Relative error between exact solution and its approximations for mesh size $30 \times 30 \times 30$

(b) Relative error between exact solution and its approximations for mesh size $100 \times 100 \times 100$

ε			1	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-10}	10^{-20}
$30 \times 30 \times 30$	Dire.	r_D	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
	Iter.	r_I	2.0	2.0	1.9	1.9	1.9	1.9	1.9	1.9
		n_I	3	3	2	2	2	2	2	2
$100 \times 100 \times 100$	Dire.	r_D	2.6	2.6	2.6	2.6	2.6	2.6	2.6	2.6
	Iter.	r_I	2.1	2.1	2.1	2.1	2.1	2.1	2.1	2.1
		n_I	3	2	2	2	2	2	2	2

(c) Ratio of computational time between resolutions of AP scheme and SP-model

Figure 8: Comparison among the 3-dimensional SP-model, the standard AP scheme (for both iterative and direct resolutions) and the SG-AP scheme in the variable ε case, where the variable ε case is taken as (4.5). The values of ε_{min}varies between 10^{-20} and 1. We compare two different mesh sizes $30 \times 30 \times 30$ and $100 \times 100 \times 100$ respectively.

5 Conclusion

In this paper, we introduced a new Asymptotic-Preserving reformulation for a highly anisotropic elliptic equation and compared it with the AP-scheme proposed initially in [10]. The new reformulation is based again on a decomposition of the unknown in its mean part and its fluctuating part, but the fluctuation equation is different. The discretization matrix associated to this new AP-reformulation is much more sparse than that of the original one, establishing thus the significant efficiency of this new method. The AP-property of the scheme is also investigated, in particular the well-posdness in the limit $\varepsilon \to 0$. Direct and iterative resolutions of the linear system are tested and compared. We note that the direct resolution is more efficient for 2D problems, however

for 3D problems the iterative resolution may be less time-consuming. Finally we consider highly heterogeneous anisotropy ratio problems, in view of real physical applications. A Scharfetter-Gummel AP-scheme is proposed in order to handle with the large gradients for variable anisotropies ε.

A One-dimensional simulations for high anisotropy ratios

In this section, we study numerically the three proposed AP-discretizations of subsection 3.3 for a one-dimensional problem extracted from the SP-model (2.2). The aim is to detect the best method in the case we have to cope with high anisotropy-gradients.

A.1 The standard AP-scheme

Let us consider the following 1-dimensional SP-model

$$(\text{SP\_1d}) \begin{cases} -\frac{\mathrm{d}}{\mathrm{d}z}\left(\frac{1}{\varepsilon(z)}\frac{\mathrm{d}u}{\mathrm{d}z}\right) + u = f, & \text{in } \Omega_z, \\ \frac{\mathrm{d}}{\mathrm{d}z}u = 0, & \text{on } \partial\Omega_z. \end{cases} \tag{A.1}$$

where u is the unknown of the problem, ε is a positive function of z and $\Omega_z = [0, L_z]$. Note that the equation (A.1) is well-posed for $\varepsilon(z) > 0$. However, taking $\varepsilon(z) := \delta\chi(z)$ and passing to the limit $\delta \to 0$, yields the degenerate problem

$$\begin{cases} -\frac{\mathrm{d}}{\mathrm{d}z}\left(\frac{1}{\chi(z)}\frac{\mathrm{d}u}{\mathrm{d}z}\right) = 0, & \text{in } \Omega_z, \\ \frac{\mathrm{d}}{\mathrm{d}z}u = 0, & \text{on } \partial\Omega_z, \end{cases} \tag{A.2}$$

which is ill-posed, as all constants are solutions.

To construct the AP-scheme corresponding to this SP-model let us decompose u into its mean part $\bar{u}$ and its fluctuating part u'. Integrating equation (A.1) in Ω_z, we get the average equation

$$\bar{u} = \bar{f}. \tag{A.3}$$

Subtracting (A.3) from (A.1), gives then the fluctuation equation

$$\begin{cases} -\frac{\mathrm{d}}{\mathrm{d}z}\left(\frac{1}{\varepsilon(z)}\frac{\mathrm{d}u'}{\mathrm{d}z}\right) + u' = f', & \text{in } \Omega_z, \\ \frac{\mathrm{d}}{\mathrm{d}z}u' = 0, & \text{on } \partial\Omega_z, \\ \overline{u'} = 0. \end{cases} \tag{A.4}$$

The system (A.3)–(A.4) is well-posed, even in the limit $\delta \to 0$. By introducing the

Lagrange multiplier λ, we obtain the following AP-reformulation

$$(\text{AP\_1d}) \begin{cases} \bar{u} = \bar{f}, \\ \displaystyle\int_0^{L_z} \left[\frac{1}{\varepsilon(z)} \frac{\mathsf{d}u'}{\mathsf{d}z} \frac{\mathsf{d}v'}{\mathsf{d}z} + u'v' \right] \mathsf{d}z + \lambda \int_0^{L_z} \frac{1}{\varepsilon(z)} v' \mathsf{d}z = \int_0^{L_z} f'v' \mathsf{d}z, \quad \forall v' \in H^1(\Omega_z), \\ \displaystyle\int_0^{L_z} u' \mathsf{d}z = 0. \end{cases} \tag{A.5}$$

This AP scheme will be discretized by a $\mathbb{P}_1$ finite element method. However, note that the difficulty here is the approximation of the high anisotropy-gradients. To face this problem, we shall propose here three different methods.

A.2 Non-Conservative AP-scheme

In this approach, we try to "break down" the high anisotropy ratio term $\frac{1}{\varepsilon}$ in the fluctuation equation (A.4). This is done, by developing $\frac{\mathsf{d}}{\mathsf{d}z}\left(\frac{1}{\varepsilon(z)}\frac{\mathsf{d}u'}{\mathsf{d}z}\right)$, in order to obtain

$$\frac{\mathsf{d}}{\mathsf{d}z}\left(\ln \varepsilon(z)\right) \frac{\mathsf{d}u'}{\mathsf{d}z} - \frac{\mathsf{d}^2 u'}{\mathsf{d}z^2} + \varepsilon u' = \varepsilon f'. \tag{A.6}$$

By injecting (A.6) in the AP-reformulation, we get another reformulation, *i.e.*

$$(\text{NC\_AP\_1d}) \begin{cases} \bar{u} = \bar{f}, \\ \displaystyle\int_0^{L_z} \left[\frac{\mathsf{d}}{\mathsf{d}z}\left(\ln \varepsilon(z)\right) \frac{\mathsf{d}u'}{\mathsf{d}z} v' + \frac{\mathsf{d}u'}{\mathsf{d}z}\frac{\mathsf{d}v'}{\mathsf{d}z} + \varepsilon u'v' \right] \mathsf{d}z + \lambda \int_0^{L_z} v' \mathsf{d}z \\ \qquad\qquad = \displaystyle\int_0^{L_z} \varepsilon f'v' \mathsf{d}z, \quad \forall v' \in H^1(\Omega_z), \\ \displaystyle\int_0^{L_z} u' \mathsf{d}z = 0. \end{cases} \tag{A.7}$$

Again, we shall discretize (A.7) by the $\mathbb{P}_1$ finite element method.

A.3 Harmonic Mean AP-scheme

The harmonic mean AP-scheme is just a special discretization of the AP reformulation (A.5). To present it, let us again consider the partition of Ω_z and the basis functions defined in section 2.5.1. Taking now in (A.5) as test functions $\kappa_k(z)$ and replacing $u'(z)$ by $\displaystyle\sum_{l=0}^{N_z+1} \alpha_l \kappa_l(z)$ gives rise for $k = 1, \cdots, N_z$ to

$$\begin{aligned} &-\frac{p_k}{\Delta z^2}\alpha_{k-1} + \left(\frac{p_k}{\Delta z^2} + \frac{p_{k+1}}{\Delta z^2}\right)\alpha_k - \frac{p_{k+1}}{\Delta z^2}\alpha_{k+1} \\ &\quad + \sum_{l=k-1}^{k+1} \alpha_l \int_{z_{k-1}}^{z_{k+1}} \kappa_l(z)\kappa_k(z)\mathsf{d}z + q_k\lambda = f_k, \end{aligned}$$

where $p_k = \int_{z_{k-1}}^{z_k} \frac{1}{\varepsilon(z)}\mathsf{d}z$, $q_k = \int_{z_{k-1}}^{z_{k+1}} \frac{1}{\varepsilon(z)}\kappa_k(z)\mathsf{d}z$, $f_k = \int_{z_{k-1}}^{z_{k+1}} f'(z)\kappa_k(z)\mathsf{d}z$. If we discretize p_k by a standard numerical quadrature formulae, the thus obtained scheme will just be a $\mathbb{P}_1$ finite element method. However, we will use instead a harmonic mean to approximate p_k, *i.e.*

$$p_k \approx \Delta z^2 \left(\int_{z_{k-1}}^{z_k} \varepsilon(z)\mathsf{d}z \right)^{-1}.$$

Finally, the integration $\int_{z_{k-1}}^{z_k} \varepsilon(z)\mathsf{d}z$ is approximated by a standard numerical quadrature, for example Gauss-Legendre quadrature. By this manner, we obtain the full harmonic mean AP scheme.

A.4 Scharfetter-Gummel AP-scheme

The Scharfetter-Gummel AP-scheme is an amelioration of the harmonic mean AP-scheme. In particular a special quadrature formulae is used to approximate p_k. Denoting $\varepsilon(z_k)$ simply by ε_k, we approximate p_k as follows

- if $\varepsilon_{k-1} \neq \varepsilon_k$,

$$\begin{aligned}
p_k &\approx \Delta z^2 \left(\int_{z_{k-1}}^{z_k} \varepsilon(z)\mathsf{d}z \right)^{-1} \\
&= \Delta z^2 \left(\int_{z_{k-1}}^{z_k} e^{\ln \varepsilon(z)}\mathsf{d}z \right)^{-1} \\
&= \Delta z^2 \left(\int_{z_{k-1}}^{z_k} \frac{\mathsf{d}e^{\ln \varepsilon(z)}}{\mathsf{d}\ln \varepsilon(z)}\mathsf{d}z \right)^{-1} \\
&\approx \Delta z^2 \left(\Delta z \frac{e^{\ln \varepsilon_k} - e^{\ln \varepsilon_{k-1}}}{\ln \varepsilon_k - \ln \varepsilon_{k-1}} \right)^{-1} \quad \text{approximate } \ln \varepsilon \text{ by } \sum_{k=0}^{N_z+1} \ln \varepsilon_k \kappa_k \\
&= \Delta z \frac{\ln \varepsilon_k - \ln \varepsilon_{k-1}}{\varepsilon_k - \varepsilon_{k-1}},
\end{aligned}$$

- if $\varepsilon_{k-1} = \varepsilon_k$,

$$p_k \approx \Delta z \frac{1}{\varepsilon_k}.$$

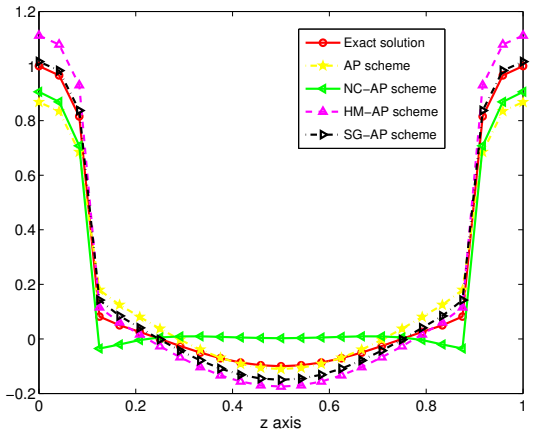
A.5 Numerical results

To compare these approaches, we take an exact solution of equation (A.1) of the form

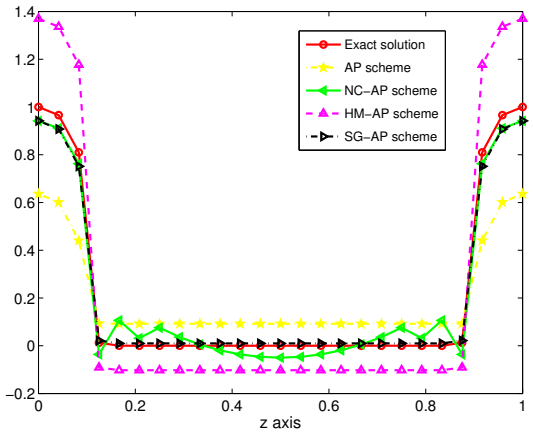
$$u_e(z) := \varepsilon(z) \cos\left(\frac{2\pi}{L_z} z\right), \tag{A.8}$$

with $\varepsilon(z)$ defined in (3.6). By injecting (A.8) into equation (A.1), we obtain the right-hand side f. We compare the SP-model, the standard AP scheme, the Non-Conservative AP scheme, the Harmonic Mean AP scheme and the Scharfetter-Gummel AP scheme respectively. For this, we choose $q = 80$, $\varepsilon_{\max} = 1$ and vary $\varepsilon_{\min}$.

Observing the numerical results in figure 9, we note that the AP scheme is robust for all choices of $\varepsilon_{\min}$, even for coarse meshes of $N_z = 25$. In figure 9(a), 9(b), we see that the SP-model can not approximate well the solution. The AP scheme as compared to the SP-model works well, but it still rather unprecise in the region near the steep gradients of the anisotropy. The NC-AP scheme exhibits oscillations near the steep interval. The HM-AP scheme is comparable to the standard AP scheme for both cases of $\varepsilon_{\min}$. Finally, the SG-AP behaves the best among all methods. The same results are also seen in table 9(c).



(a) $\varepsilon_{\min} = 10^{-1}$



(b) $\varepsilon_{\min} = 10^{-10}$

$\varepsilon_{\min}$	SP-model	AP scheme	NC-AP scheme	HM-AP scheme	SG-AP scheme
10^{-1}	26.4022	0.1652	0.1486	0.1532	0.0734
10^{-10}	11.5878	0.4310	0.1171	0.4434	0.0653

(c) Relative errors between the exact solution and the approximate ones

Figure 9: Comparison between SP-model, AP scheme, Non-Conservative AP scheme, Harmonic Mean AP scheme and Scharfetter-Gummel AP scheme. We take $N_z = 25$. (a) Plots of the approximate solutions via the different methods in the case $\varepsilon_{\min} = 10^{-1}$; (b) Same plots in the case $\varepsilon_{\min} = 10^{-10}$; (c) Relative errors between the exact solution and the approximate ones for both cases $\varepsilon_{\min} = 10^{-1}$ and $\varepsilon_{\min} = 10^{-10}$.

References

[1] S. F. Ashby, R. D. Falgout, T. W. Fogwell, A. F. B. Tompson, *A numerical solution of groundwater flow and contaminant transport on the CRAY T3D and C90 supercomputers*, Int. J. High Perform. Comp. Appl., Vol. 13 (1999), pp 80–93.

[2] W. G. Baker, *Electric Currents in the Ionosphere. II. The Atmospheric Dynamo*, Phil. Trans. R. Soc. Lond. A, Vol. 246,(1953) pp 295–305.

[3] C. Besse, J. Claudel, P. Degond, F. Deluzet, G. Gallice, C. Tessieras, *Numerical simulations of the ionospheric striation model in a non-uniform magnetic field*, Comp. Phys. Comm., Vol. 176, No. 2 (2007) pp 75-90.

[4] C. Besse, J. Claudel, P. Degond, F. Deluzet, G. Gallice, C. Tessieras, *A model Hierarchy for Ionospheric Plasma modeling*, Math. models Methods Appl. Sci, Vol. 14, No. 3 (2004), pp 393–415.

[5] C. Besse, P. Degond, HJ. Hwang, R. Poncet, *Nonlinear Instability of the Two-Dimensional Striation Model About Smooth Steady States*, Communications in Partial Differential Equations, Vol. 32, Issue 7 (2007), pp 1017-1041.

[6] C. Besse, F. Deluzet, C. Yang, *Numerical simulations of the ionospheric dynamo model in non-uniform magnetic field*, in preparation.

[7] D. Biliza, *International reference ionosphere 2000*, Radio Sci., Vol. 36, No. 2 (2001), pp 261–275.

[8] P. Degond, F. Deluzet, L. Navoret, A.-B. Sun, M.-H. Vignal, *Asymptotic-Preserving Particle-In-Cell method for the Vlasov-Poisson system near quasineutrality*, J. Comput. Phys, 229 (2010), pp 5630-5652.

[9] P. Degond, F. Deluzet, D. Maldarella, J. Narski, C. Negulescu and M. Parisot, *Hybrid model for the coupling of an symptotic preserving scheme with the asymptotic limit mode: the one dimensional case.*, submitted.

[10] P. Degond, F. Deluzet, C. Negulescu, *An asymptotic preserving scheme for strongly anisotropic elliptic problems* , SIAM-MMS (Multiscale Modeling and Simulation), Vol. 8, No. 2 (2010), pp 645–666.

[11] P. Degond, A. Lozinki, J. Narski, C. Negulescu, *An asymptotic preserving method for highly anisotropic elliptic equations based on a micro-macro decomposition*, submitted.

[12] J. Du, R. J. Stening, *Simulating the ionospheric dynamo – II. Equatorial electric fields*, Journal of Atmospheric and Solar-Terrestrial Physics, Vol. 61, Issue 12 (1999), pp 925–940.

[13] E. C. Gartland, *On the uniform convergence of the Scharfetter-Gummel discretization in one dimension*, SIAM J. Numer. Anal. Vol. 30, No. 3 (1993), pp 749–758.

[14] John Keith Hargreaves. *The solar-terrestrial environment: an introduction to geospace–the science of the terrestrial upper atmosphere, ionosphere, and magnetosphere*. Cambridge University Press, 1992.

[15] N. J. Higham, *FORTRAN Codes for Estimating the One-Norm of a Real or Complex Matrix with Applications to Condition Estimation*, ACM Transaction on Mathematical Software, Vol. 14, No. 4 (1988), pp 381–396.

[16] N. J. Higham, F. Tisseur, *A Block Algorithm for Matrix 1-Norm Estimation, with an Application to 1-Norm Pseudospectra*, SIAM J. Matrix Anal. Appl. Vol. 21, No. 4 (2000), pp 1185–1201.

[17] T. Y. Hou, X. H. Wu, *A Multiscale Finite Element Method for Elliptic Problems in Composite Materials and Porous Media*, J. Comput Phys. , 134, (1997), pp 169–189.

[18] Huba, J.D., G. Joyce, and J.A. Fedder, *Sami2 is Another Model of the Ionosphere (SAMI2): A new low-latitude ionosphere model*, J. Geophys. Res., Vol. 105, No. 23, 035, (2000).

[19] Huba, J.D., G. Joyce, and J.A. Fedder, *Simulation study of mid-latitude ionosphere fluctuations observed at Millstone Hill*, Geophys. Res. Lett., Vol. 30, No. 18, (2003), pp 1943.

[20] R. D. Hunsucker and J. K. Hargreaves. *The High-Latitude Ionosphere and its Effects on Radio Propagation*. Cambridge Atmospheric and Space Science Series. Cambridge University Press, 2002.

[21] S. Jin, *Efficient Asymptotic-Preserving (AP) schemes for some multiscale kinetic equations*, SIAM J. Sci. Comp., Vol. 21 (1999), pp 441–454.

[22] K. Kawano-Sasaki, S. Miyahara, *A study on three-dimensional structures of the ionospheric dynamo currents induced by the neutral winds simulated by the Kyushu-GCM*, Journal of Atmospheric and Solar-Terrestrial Physics, Vol. 70, Issues 11-12 (2008), pp 1549–1562.

[23] M. C. Kelley, W. E. Swartz, J.J. Makela, *Mid-Lattitude ionospheric fluctuation spectra due to secondary $E \times B$ instabilities*, J. Atmos. Solar-Terr. Phys., Vol. 66 (2004), pp 1559–1565.

[24] M. J. Keskinen, S. L. Ossakow, B. G. Fejer, *Three-dimensional nonlinear evolution of equatorial ionospheric spread-F bubbles*,Geophys. Res. Lett., Vol. 30 (2003), pp 41–44.

[25] T. Manku, A. Nathan, *Electrical properties of silicon under nonuniform stress*, J. Appl. Phys. **74** (1993), pp 1832.

[26] J. H. Piddington, *Irregularities in the upper ionosphere*, Planetary and Space Science, Vol. 12, Issue 2 (1964), pp 127–136.

[27] Henry Rishbeth and Owen K. Garriott. *Introduction to ionospheric physics [by] Henry Rishbeth [and] Owen K. Garriott.* International geophysics series ; v. 14. Academic Press, New York,, 1969. Revision and expansion of Stanford Electronics Laboratories report SU-SEL-64-111 issued in 1964 under title: Introduction to the ionosphere and geomagnetism. Bibliography: p. 275-309.

[28] O. Schenk and K. Gartner, *Solving Unsymmetric Sparse Systems of Linear Equations with PARDISO*, Journal of Future Generation Computer Systems, Vol. 20, No. 3 (2004), pp 475–487.

[29] O. Schenk, A. Waechter, and M. Hagemann, *Matching-based Preprocessing Algorithms to the Solution of Saddle-Point Problems in Large-Scale Nonconvex Interior-Point Optimization.* Journal of Computational Optimization and Applications, Vol. 36, No. 2-3(2007), pp 321–341.

[30] A. Singh, K. D. Cole, *A numerical model of the ionospheric dynamo – I. Formulation and numerical technique*, Journal of Atmospheric and Terrestrial Physics, Vol. 49, Issue 6 (1987), pp 521–527.

[31] N. Saito, *An interpretation of the Scharfetter-Gummel finite difference scheme*, Proc. Japan Acad., Vol. 82, Ser. A (2006), pp 187–191.

[32] D. L. Scharfetter, H. K. Gummel, *Large-signal analysis of a silicon read diode oscillator*, IEEE Trans. Electron Devices ED-16 (1969), pp 64–77.

[33] A.-M. Tréguier, *Modélisation numérique pour l'océanographie physique*, Ann. math. Blaise Pascal, tome 9, no. 2 (2002), pp 345–361.

[34] T. Tsuchiya, K.Yoshida and S. Ishioka, *Yamamoto's principle and its applications to precise finite element error analysis*, J.comput. Appl. Math. Vol. 152, No. 1-2 (2003), pp 507–532.

[35] H. A. Van Der Vorst, *Krylov subspace iteration*, Computing in science & engineering, (2000), pp 32–37.

[36] W.-W. Wang, X.-C. Feng, *Anisotropic diffusion with nonlinear structure tensor*, Multiscale Model Sim., Vol. 7, no. 2 (2008), pp 963–977.

[37] X. Wei, *Three Dimensional Rigorous Model for Optical Scattering Problems*, PhD thesis, Technische Universiteit Delft, 2006, pp 73-78.

[38] J. Weickert, *Anisotropic Diffusion in Image Processing* , Teubner, Stuttgart, (1998).